ARTS AND SCIENCE
ARTS AND SCIENCE PRESS PTE. LTD

Research Article

Earthquake prediction using artificial intelligence in the Ferghana depression (Uzbekistan)

Ikram Atabekov[1,2], Aziz Atabekov[3], Rustam Lutfullaev[1], Azizbek Biimbetov[3], Jasur Mamarakhimov[1,3]

[1] *Institute of Seismology, Zulfiyakhanum Street 3, 100128 Tashkent, Uzbekistan*

[2] *Institute of Biophysics and Biochemistry at National University of Uzbekistan, University str., 4, 100174 Tashkent, Uzbekistan.*

[3] *Digital Technologies and Artificial Intelligence Development Research Institute, Buz-2, 17A, 100125, Tashkent, Uzbekistan*

* Corresponding author: Ikram Atabekov, atabekovi@mail.ru, ORCID ID: 0000-0002-6500-1114, tel.+998909151373

ABSTRACT

Earthquake prediction remains a formidable challenge due to its theoretical and practical complexities. The multifactorial nature of earthquakes leads to diverse anomalies, which are potential precursors. However, the intricate earthquake process and limited knowledge of the Earth's crust structure restrict the accuracy of these predictions. This study introduces an advancements using machine learning and deep learning methods, notably the Kora 3 and Kora 4 algorithms, to identify key earthquake features. We employed LSTM and RNN deep learning algorithms to predict earthquakes of magnitude M≤4.3 without temporal data. Our methodology was applied to the 2022 earthquake monitoring in the Fergana depression, demonstrating significant advancements in seismic event prediction.

Keywords: geodynamical modeling; stress; earthquake; machine learning; deep learning

1. Introduction

The Fergana depression, located in Central Asia's northeastern sector, is an intermountain basin encircled by significant mountain ranges: the Kuramin and Chatkal Ranges to the north and northwest, the Fergana Range to the east and northeast, and the Alay and Turkestan Ranges to the south. This valley connects to the Turan lowland through a narrow western passage. The topography of the Fergana depression is marked by contrasting elevations: the surrounding mountains soar to heights of up to 5,000 meters in the south, around 4,000 meters in the east, and between 2,000 to 3,000 meters in the north and northwest. At its foothills, the landscape is characterized by lower mountain ranges, adyrs, and intermountain valleys. The central area features a pluvial-alluvial plain, irrigated by the region's largest river, the Sirdarya, and its principal tributaries, the Narin and Qaradarya (refer to **Figure 1**).
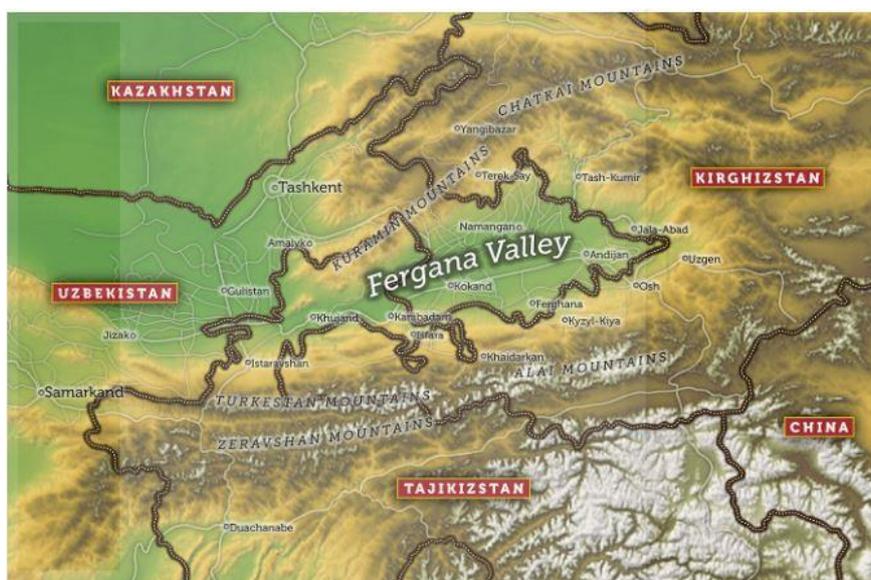
Figure 1. Ferghana depression (Uzbekistan) and adjacent countries.

The Fergana depression in Uzbekistan is known for its seismic activity, frequently experiencing moderate earthquakes. Historical records indicate that this region has witnessed eight significant seismic events, each with a magnitude of M≥6.5, as detailed in **Table 1**.

Table 1. Eight significant seismic events occurred in the Fergana area of Uzbekistan.

Year	Month	Day	H	M	S	Lat	Lon	Depth	Mag
1896	0	0	0	0	0	41,5	70,9	20	6,9
1902	0	0	0	0	0	40,8	72,3	20	6,6
1923	0	0	0	0	0	39,59	69,19	20	6,6
1924	0	0	0	0	0	40,5	73,1	20	6,6
1924	0	0	0	0	0	40,59	73,19	20	6,7
1937	0	0	0	0	0	42,1	70,9	20	6,7
1946	0	0	0	0	0	41,9	72	20	7,9
1977	1	31	14	26	14	40,11	70,79	20	6,6

Current research in seismic studies extensively employs various artificial intelligence algorithms to predict earthquake magnitudes from seismic time series. These methodologies, utilizing diverse earthquake features, have been successful globally[1-14]. These predictive models generally forecast magnitudes without being region-specific. An exception is found in a study[15] that divides the area under consideration into eight zones, attempting to associate earthquake magnitudes with specific zones. However, such research is scarce for Central Asia. The notable exception is the pioneering work by the Institute of Physics of the Earth, which employed a basic pattern recognition algorithm for this region[16]. This approach was later refined by scientists from the Institute of Earth Physics of the Russian Academy of Sciences and the Institute of Earthquake Forecast Theory and Mathematical Geophysics of the Russian Academy of Sciences. They enhanced algorithms for pinpointing locations of significant earthquakes and successfully implemented these in various global regions. A comprehensive review of this research and its applications can be found in[17-19].

The primary indicators of earthquake susceptibility are morphometric features, particularly at points where faults intersect. In a study focusing on the Fergana depression[16], only four epicenters of significant earthquakes (magnitude M≥6.5) were identified. Subsequent analysis suggested the possibility of a major

earthquake occurring at one specific node in the depression. Yet, to date, no earthquakes with a magnitude of M≥6.5 have been recorded at this node. This leads us to believe that relying solely on morphometric features may be insufficient for accurate earthquake prediction. These features primarily represent the Earth's crust deformation, but as observed in viscoelastic materials, significant deformations do not necessarily indicate an imminent fault.

The seismic process, as noted by Gzovsky[20], is primarily governed by the maximum shear stress and the highest gradients of movement velocities. He observed that all-around compression tends to inhibit earthquake occurrences. Consequently, morphological features serve as indicators of the Earth's crust's involvement in the deformation process. A key predictor of an impending earthquake is the assessment of energy accumulation or the proximity of tangential stress to the crust's ultimate local strength. However, Gzovsky identified inconsistencies in using geomorphological features from the Quaternary period for predicting strong earthquakes through pattern recognition. These features, based on average movement velocity gradients over extended periods, fail to capture the dynamics of these movements. He argued that since the current seismic stress state and the one during the formation of crustal faults can vary significantly, fault activity should be considered a crucial criterion. Information about tectonic activity can be garnered through seismotectonics, geodesy, and seismic methods. Riznichenko[21] emphasized assessing earthquake potential by analyzing seismotectonic flow of rock masses, quantifiable through mathematical modeling. Crustal deformations offer valuable insights into the evolution of seismic processes and potential future earthquakes. In this paper, we enhance prediction accuracy by utilizing not only seismological time series and morphostructural features but also other critical characteristics like stress state and recent movements, as determined through numerical modeling[22]. In our study, Earth's crust deformations and stresses in the Fergana depression, specifically at depths of 15-20 km, were modeled using Stokes equations, considering the primary active faults. The internal stresses for Central Asia from[23] were applied as boundary conditions. Seismological data indicate that eight significant earthquakes (M≥6.5) occurred at approximately 20 km depth (**Table 1**). For effective earthquake monitoring, it's crucial to consider variations in the stress state over preceding time intervals. Changes in the stress state following earthquakes are determined through modeling the earthquake mechanism as a double dipole without a moment[24-25].

2. Methodology

The identification of strong earthquake locations in this study is achieved through the application several methodologies, specifically Cora and Random Forest, and machine learning and deep learning architectures such as ANN (Artificial Neural Network) and LSTM (Long Short-Term Memory). The core principle of pattern recognition employed here involves spatial division via a hyperplane. This hyperplane is defined by a vector comprising a set of event attributes, effectively classifying the data into two distinct groups. Various algorithms differ in their approach to determining the optimal hyperplane, each guided by unique criteria. The majority of machine learning techniques in earthquake prediction are adaptations of these algorithms, tailored to address this specific challenge.

The foundational principle of the Cora 3 algorithm involves compiling a dataset with predictive features that characterize earthquakes. To construct these datasets, the region of interest, specifically the Fergana Depression, was segmented into 17 distinct zones (see **Figure 2**). This segmentation strategy was grounded in the amalgamation of individual zones from the computational grid used in the numerical modeling of the Earth's crust's stressed state in the Fergana Depression[23]. This approach also aimed to avoid excessively narrow zoning, which could limit the representativeness of earthquakes for determining critical seismic parameters such as Richter-Gutenberg coefficients, earthquake counts, and other related seismic metrics.
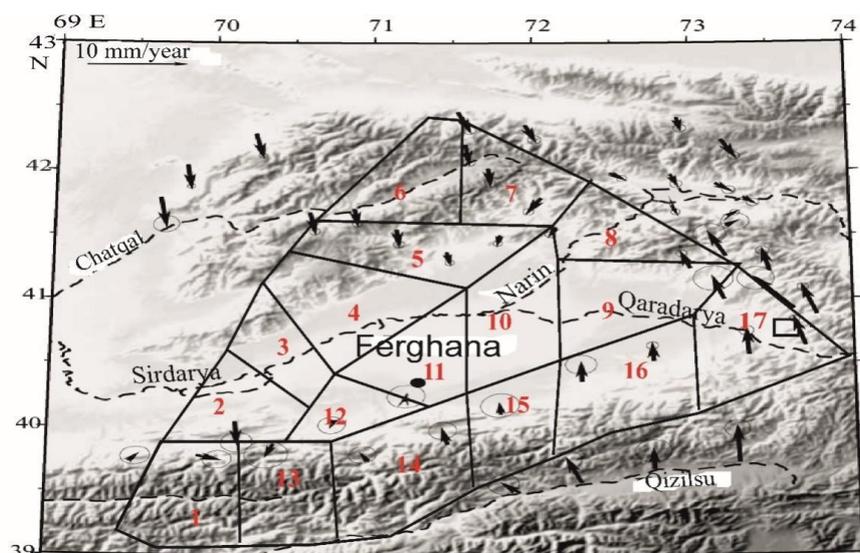
Figure 2. Ferghana depression divided into zones.

In each of the 17 zones, the following specific features were determined (Table 2):

1) Distance to the nearest mountainous country R≤ 30 km;

2) 30<R≤ 70;

3) R>70;

4) Morphological combinations: mountains with foothills and foothills;

5) mountains with foothill plains;

6) mountains with other mountains;

7) foothills with foothill plains;

8) relief height h≤1 km;

9) 1<h≤ 2 km;

10) h>2 km;

11) height difference dh≤ 1.5 km;

12) 1.5<dh≤ 2.5;

13) dh>2.5 km;

14) number of faults 1≤n≤2;

15) n>2;

16) crossing faults;

17) length of the main fault L≤300 km;

18) 300<L≤700 km;

19) L>700 km;

20). Relative tangential stress versus instantaneous shear strength τ>0.9;

21) Modern movements: upward vertical velocities;

4

22) Modern movements: down vertical velocities;

23) Nature of the stress state: regional compression;

24) Regional stretching;

25) Nature of the earthquake: fault;

26) Reverse;

27) Shift;

28) The difference between the Richter-Guttenberg coefficient and the average value of the entire territory of the Fergana depression $\Delta b \geq 0$;

29) The difference in the relative activity of earthquakes from the average value of the entire territory $\Delta A10 \geq 0$;

30) The presence of earthquakes with magnitude $M \geq 6.5$.

Table 2 is structured with the feature numbers displayed horizontally and the zone numbers vertically. A value of 1 is assigned if the specific feature is present in a zone, and 0 if it is absent.

Table 2. Earthquake features(horizontally) and zones (in vertical).

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
1	1	0	0	0	0	0	0	1	0	0	1	0	0	1	1	0	1	0	0	0	1	0	1	0	1	1	1	0	0	0
2	0	1	0	0	0	0	1	1	0	0	0	1	0	1	0	1	0	0	1	1	0	1	0	1	1	0	1	0	0	1
3	1	0	0	1	0	0	1	1	0	0	1	0	0	0	1	1	0	0	1	0	1	0	0	1	1	0	0	0	0	0
4	1	0	0	1	0	0	1	1	0	0	1	0	0	0	1	0	0	1	0	0	1	0	1	0	1	0	0	1	0	0
5	1	0	0	1	0	1	0	1	0	0	0	0	1	0	1	1	0	0	1	1	1	0	1	0	0	1	0	0	0	1
6	1	0	0	0	0	0	0	1	0	0	1	0	0	0	1	1	0	1	1	1	1	0	1	0	1	1	0	0	0	1
7	1	0	0	1	0	0	0	0	0	1	0	1	0	1	0	1	0	1	0	1	1	1	0	1	0	0	0	1	0	1
8	1	0	0	0	1	0	0	1	0	0	1	1	0	0	1	0	0	0	1	0	1	0	1	0	1	0	0	0	0	0
9	1	0	0	0	1	0	0	1	0	0	0	1	0	0	1	0	0	0	1	1	0	1	1	0	1	1	0	1	0	1
10	0	0	1	0	0	0	1	1	0	0	1	0	0	0	1	0	0	0	1	1	0	1	1	0	1	0	0	1	0	1
11	0	0	1	0	0	0	1	1	0	0	1	0	0	0	1	1	0	1	0	1	0	1	1	0	1	0	0	0	0	1
12	0	0	1	0	0	0	1	1	0	0	1	0	0	1	0	0	1	0	0	0	0	1	0	1	1	0	0	0	0	0
13	1	0	0	0	1	1	0	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0	1	0	0	1	0	0	0	0
14	1	0	0	0	1	1	0	0	0	0	1	0	0	0	1	0	1	0	0	1	1	0	1	0	0	1	0	1	1	0
15	1	0	0	0	1	1	0	0	0	0	1	0	0	0	1	0	0	0	1	0	1	0	1	0	0	1	0	1	0	0
16	1	0	0	0	1	1	0	0	0	0	1	0	0	0	1	0	0	0	1	0	1	0	1	0	1	1	0	1	0	0
17	1	0	0	0	0	1	1	0	0	1	1	0	0	0	1	0	0	0	1	0	1	0	1	0	1	0	0	1	1	0

Features 1-13 were identified using an online topographic map. Presently, there exists no contemporary, universally accepted fault map of the Fergana intermountain basin; these maps often vary based on the tectonic interpretations of different researchers and are not directly discernible through geophysical methods. The location and number of faults (**Features 14-19**) were determined based on the findings from our previous studies[22], as illustrated in **Figure 3**.
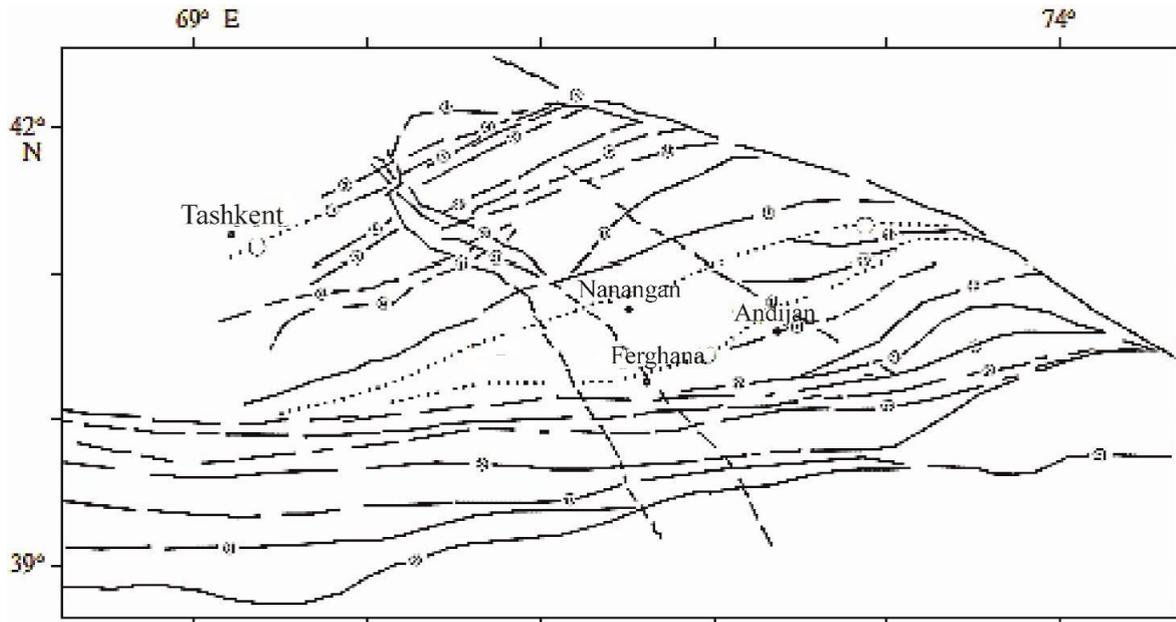
Figure 3. Active faults of the Ferghana depression.

Modern movements (**Features 21-22**), stresses and geodynamic states (**Features 23-27**) were determined from the modeling results[22] (**Figures 4-6**).
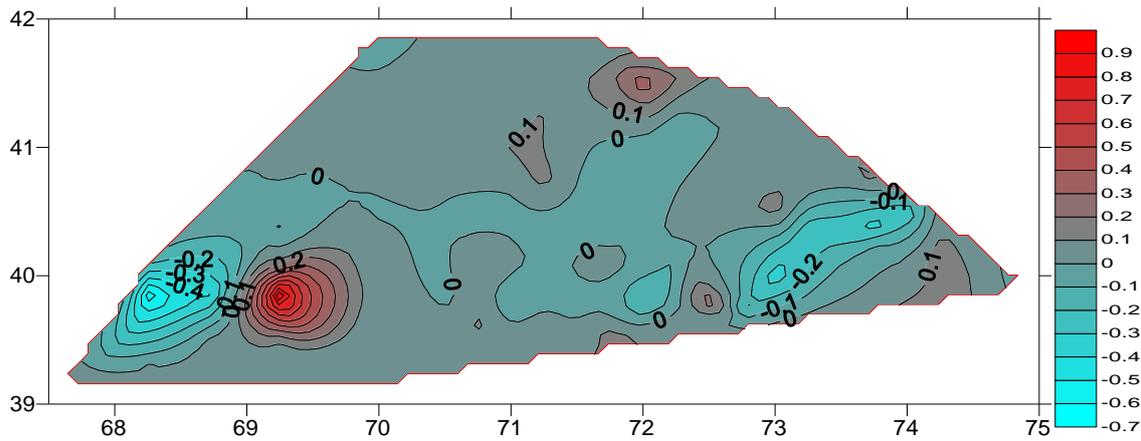

Figure 4. Vertical velocities of the Earth's crust in the Fergana depression.
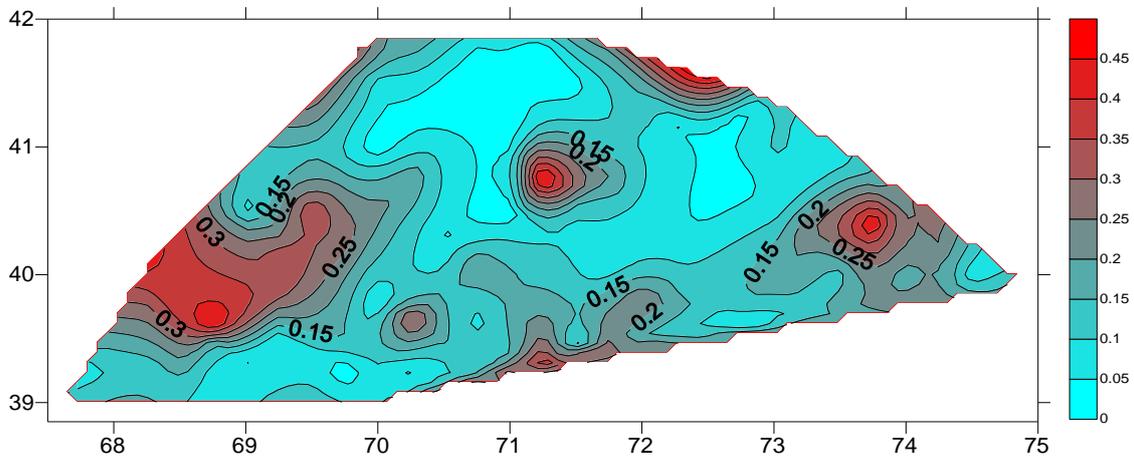

Figure 5. Tangential stresses of the Earth's crust in the Ferghana depression.
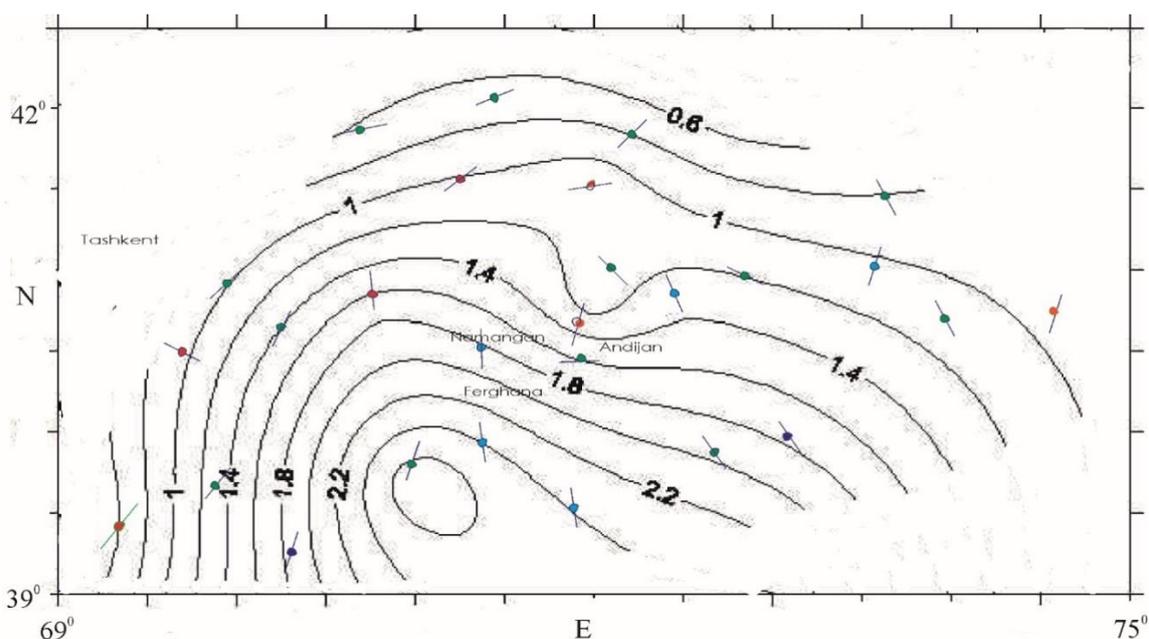
Figure 6. The highest horizontal stress (compression) σ1 (in 1:102 MPa) for depths of 15-20 km.

Horizontal is eastern longitude, vertical is northern latitude. The orientation of the arrows indicates the direction of σ1. Anderson's classification of the geodynamic state in the event of an earthquake indicates the possibility of a reverse fault (red), a normal fault (blue), and displacement (green).

To address the challenge of earthquake prediction, in accordance with the Cora 3 algorithm, a training table is generated with each row representing a description of seismic zones, denoted as $A = a_{ij} (i=1..n, j=1..m)$. All zones are divided into two classes, characterized by the presence or absence of predicted seismic manifestations - earthquakes of magnitude $M \geq M_0$. For each class, characteristic sets of features are found. As such sets of features, the algorithm considers all possible sets of 3 features from among the given n features. The characteristic set of a class consists of all sets of feature values that occur only in a given class and do not occur in another class.

Thus, certain sets of sets of feature values $\{a_r, a_u, a_v\}$ are called characteristics of class K_1. The characteristics $\{a_r, a_u, a_v\}$ of class K_2 are determined similarly. At the stage of recognition of a given object $S = (a_1, a_2, …, a_n)$, its sets of feature values are compared with the characteristics of the corresponding triplets for classes K_1 and K_2. Let us denote $G(S, K_i)$ – a measure of the proximity of an object S to class K_i - the number of sets of feature values S that coincide with the characteristics of class K_i. , $i=1,2$. If $G(S, K_1) > G(S, K_2) + \Delta$, where $\Delta > 0$, then object S is assigned to class K_1; otherwise, it is assigned to class K_2. The quality of the recognition algorithm is determined by the proportion of correctly recognized objects.

The Core 3 algorithm was created early in the computer revolution and was designed for low-power machines. Taking into account the realities of today, we decided to slightly complicate the solution of the problem so that the characteristic classes K_1 and K_2 contain not three but four attributes $\{a_r, a_u, a_v, a_w\}$. With the same signs (table), the Cora 4 algorithm was unable to determine the presence of an earthquake in zone 6 (table). However, it anticipates a major earthquake in Zone 12, similar to Cora 3, in the future. Naturally, the requirements for characterizing earthquakes become more stringent. Dividing the territory under consideration into smaller zones would make it possible to more accurately determine some characteristics, especially morphometric ones. However, this approach leads to a reduced number of earthquakes being associated with

7

specific zones, raising concerns about the representativeness of these earthquakes for identifying additional characteristics.

In contrast to other machine learning algorithms, LSTM (long short-term memory) features a mechanism for storing information in memory that, at its most basic level, relies on the previous state h_{t-1}. It takes the current input x_t and performs calculations to determine which of them to store in memory as the hidden state h_t.

The architecture (**Figure 7**) and the system of equations that implements this architecture have a similar form of[13]:

$$i_t = \sigma \ (W_{ix}x_t + W_{ih}h_{t-1} + W_{ic}c_{t-1} + b_i)$$

$$f_t = \sigma \ (W_{fx}x_t + W_{fh}h_{t-1} + W_{fc}c_{t-1} + b_f)$$

$$c_t = f_t \circ c_{t-1} + i_t \circ (W_{cx}x_t + W_{ch}h_{t-1} + b_c) \quad (1)$$

$$o_t = \sigma \ (W_{ox}x_t + W_{oh}h_{t-1} + W_{oc}c_t + b_o)$$

$$h_t = o_t \circ \phi(c_t)$$

Here, *i, f, o, c,* and h denote the input gate, forget gate, output gate, cell state, and hidden vector, respectively. σ is a sigmoid function that takes values [-1; 1]. W_{ic}, W_{fc} and W_{oc} are the peephole connection matrices, which connect the cell state to the input gate, forget gate, and output gate, respectively. Similarly, W_{ix}; W_{fx}; W_{ox} and W_{cx} are the weight matrices connecting the input vector x_t and the input gate, forget gate, output gate and cell state, respectively. The icon ○ is a Hadamard product that is a binary operation on two matrices of the same dimension, the result of which is a matrix of the same dimension, in which each element with indices *i and j* is the product of elements with indices *i and j* of the original matrices.
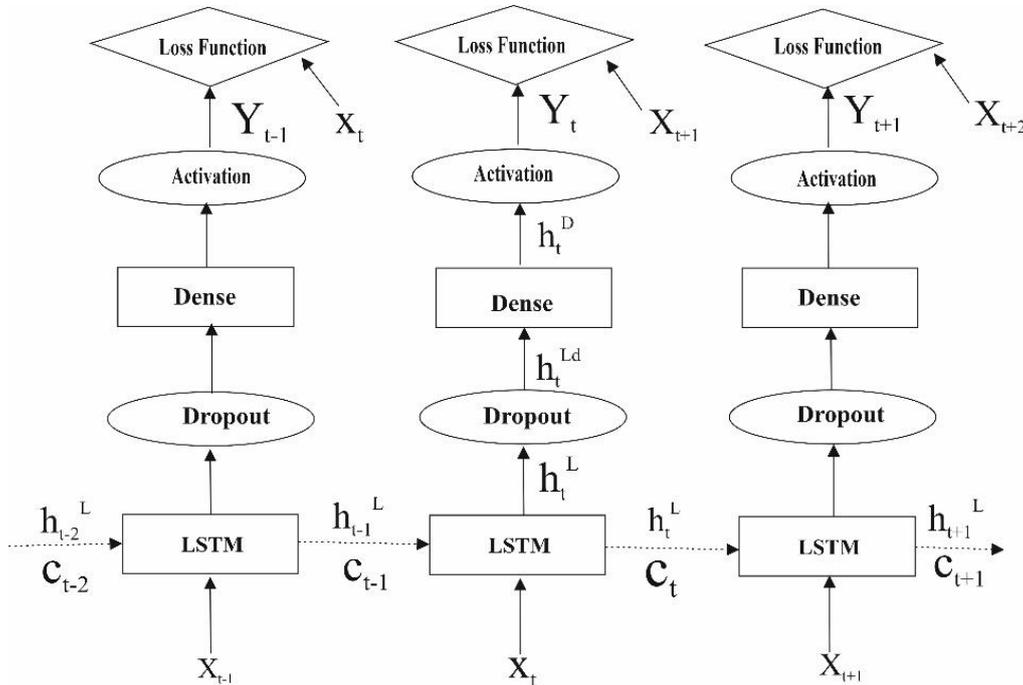

Figure 7. Architecture LSMT built to solve the problem.

During the pruning process of the neural network, less significant or redundant weights are strategically removed. This reduction aims to simplify the model's architecture and decrease its size, while still preserving its operational efficacy. The loss function plays a crucial role in quantifying the divergence between predicted

outputs and actual values. It serves as a critical tool during the neural network's training phase, allowing for the evaluation of the model's performance and guiding adjustments in its structure or parameters. Typically established in the cross-validation phase, the data set is partitioned into training (80%) and testing (20%) subsets, facilitating the optimization of the loss function.

3. Analysis of model solutions

Table 3 shows the testing results. Earthquakes 0 and 1 indicate the values in column 30 of Table 1. Prediction 0 or 1 means recognition match. The numerator and denominator of the assessment indicate the share of $G(S, K_1)/G(S, K_2)$ in classes K_1 and K_2. Correct or incorrect seismicity prediction is marked with TRUE or FALSE. In this scenario, 16 out of 17 events were accurately identified, resulting in a recognition accuracy of 94.1%. Meanwhile, the algorithm indicates that in zone 12, the likelihood of a strong earthquake ($M≥6.5$) is comparatively high, with a ratio of 169.7 against 66.78.

Table 3. Testing by Cora 3 algorithm.

Zones	Class K_1
1	Earthquake 0 forecast 0 value: 98.00/45.14 (TRUE)
3	Earthquake 0 forecast 0 value: 69.44/68.71 (TRUE)
4	Earthquake 0 forecast 0 value: 82.00/53.43 (TRUE)
8	Earthquake 0 forecast 0 value: 60.11/53.43 (TRUE)
12	***Earthquake 0 forecast 1 value: 66.78/169.7 (FALSE)***
13	Earthquake 0 forecast 0 value: 132.56/8.14 (TRUE)
14	Earthquake 0 forecast 0 value: 126.56/45.29 (TRUE)
15	Earthquake 0 forecast 0 value: 119.44/3.00 (TRUE)
16	Earthquake 0 forecast 0 value: 110.89/1.14 (TRUE)
17	Earthquake 0 forecast 0 value: 102.56/33.86 (TRUE)
	Class K_2
2	Earthquake 1 forecast 1 value: 73.70/133.67 (TRUE)
5	Earthquake 1 forecast 1 value: 50.80/94.83 (TRUE)
6	Earthquake 1 forecast 1 value: 37.20/65.17 (TRUE)
7	Earthquake 1 forecast 1 value: 71.90/117.33 (TRUE)
9	Earthquake 1 forecast 1 value: 55.00/118.00 (TRUE)
10	Earthquake 1 forecast 1 value: 32.10/125.17 (TRUE)
11	Earthquake 1 forecast 1 value: 39.80/133.50 (TRUE)

Using the same features outlined in **Table 2**, the Cora 4 algorithm was not able to identify an earthquake occurrence in zone 6, as shown in **Table 4**. Nevertheless, it predicts a significant earthquake in Zone 12 in the future, aligning with the Cora 3 algorithm's findings. However, the true positive rate of recognition quality diminishes to 88.2%. This necessitates more stringent criteria for earthquake characterization. Subdividing the area into smaller zones could enhance the accuracy in identifying certain features, particularly morphometric ones. Yet, this leads to a decrease in the number of earthquakes associated with specific zones, raising concerns about the representativeness of these earthquakes for identifying other characteristics.

Table 4. Testing by Cora 4 algorithm.

Zones	Class K_1
1	Earthquake 0 forecast 0 value: 734.78/122.00 (TRUE)
3	Earthquake 0 forecast 0 value: 584.89/347.86 (TRUE)
4	Earthquake 0 forecast 0 value: 738.22/229.43(TRUE)
8	Earthquake 0 forecast 0 value: 567.33/343.43 (TRUE)
12	*Earthquake 0 forecast 1 value: 406.22/884.71 (FALSE)*
13	Earthquake 0 forecast 0 value: 1217.00/6.29(TRUE)
14	Earthquake 0 forecast 0 value: 1039.11/189.00 (TRUE)
15	Earthquake 0 forecast 0 value: 1164.22/7.29(TRUE)
16	Earthquake 0 forecast 0 value: 1071.44/2.43 (TRUE)
17	Earthquake 0 forecast 0 value: 860.67/124.14 (TRUE)
	Class K_2
2	Earthquake 1 forecast 1 value: 433.10/541.17 (TRUE)
5	Earthquake 1 forecast 1 value: 388.10/474.83 (TRUE)
6	*Earthquake 1 forecast 1 value: 432.70/322.33 (TRUE)*
7	Earthquake 1 forecast 1 value: 376.50/496.17 (TRUE)
9	Earthquake 1 forecast 1 value: 451.50/545.33 (TRUE)
10	Earthquake 1 forecast 1 value: 319.30/671.17 (TRUE)
11	Earthquake 1 forecast 1 value: 303.10/701.67 (TRUE)

The Random Forest algorithm, a versatile machine learning tool, is employed for classification, regression, and clustering tasks[27]. Its core strategy involves the integration of a substantial ensemble of decision trees. Individually, these trees may yield low classification accuracy, but collectively, they produce robust results. Decision-making is based on the majority vote among these trees, each constructed independently. The strength of this ensemble approach lies in the diversity and effectiveness of its base algorithms, with each tree developed on a unique training sample and incorporating randomness in split decisions. In our study, the algorithm was trained on all zones except one, which was then used for the recognition process. This application achieved a 100% accuracy rate, correctly identifying zones with earthquakes of magnitude M≥6.5, as well as those without, as indicated in **Table 5**. The weights generated from this algorithm implementation are depicted in **Figure 8**.

Table 5. Results by Random forest algorithm.

Zones	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Real	0	1	0	0	1	1	1	0	0	0	0	1	1	1	1	1	1
Test	0	1	0	0	1	1	1	0	0	0	0	1	1	1	1	1	1

It is evident that the various indicators used in earthquake prediction are not of equal significance. Understanding the importance of each feature can shed light on the mechanisms underlying seismic activities and aid in developing a more accurate seismic model. As expected, the proximity of maximum shear stress to the crust's instantaneous strength limit is a leading indicator[20]. This is followed by the presence of active fault intersections[16] and current movements, particularly vertical velocities [21]. Other significant factors include the length of the main fault[18], as well as morphometric features such as small elevation changes[11], among others. In terms of fault types, reverse faults [26] and strike-slip faults[27] are more prevalent than normal faults[25].

The region exhibits more expansion[24] than compression[23], a characteristic typical of rift zones. Interestingly, the Richter-Guttenberg coefficient[28] and earthquake activity[29] carry lesser weights in the model. This could be attributed to the reduced representativeness of earthquakes in more narrowly defined zones.
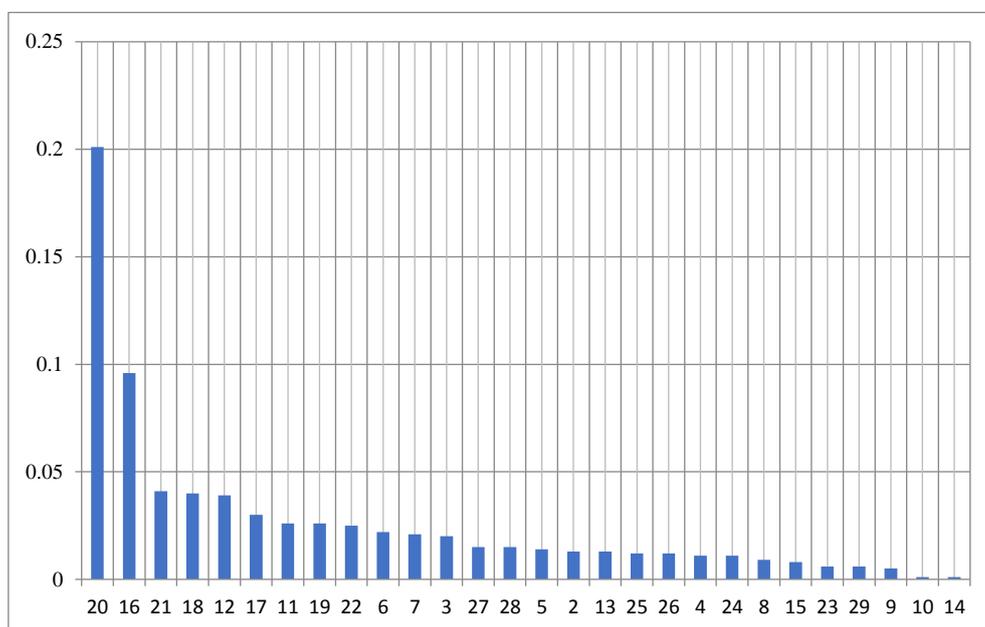

Figure 8. The feature weights are shown in Table 2 as a result of the implementation of the random forest algorithm.

The study's results demonstrate that machine learning algorithms can effectively identify key features for monitoring significant earthquakes and pinpoint potential locations of such events, although they do not specify the timing of their occurrence.

The same set of features was employed in predicting earthquakes using deep learning techniques. Two distinct architectures were developed: one based on an Artificial Neural Network (ANN) and the other on a recurrent Long Short-Term Memory (LSTM) framework comprising 128 modules. The input for these models includes seismological data specific to each zone, such as earthquake dates, geographic coordinates, source depth, earthquake magnitude, and an additional 40 attributes listed in the table. The network training utilized seismological data spanning several years. Model accuracy at each layer was assessed using the root mean square error (RMSE) method, and optimization of weights was performed using the Adam optimization algorithm. In the ANN-based model, the output layer predicts a single parameter—the magnitude of the next earthquake. Conversely, the LSTM-based model outputs two parameters: the magnitude and zone of the next earthquake. The RMSE for the ANN model is 0.12 and 0.14 for the LSTM model, as detailed in **Table 6**.

Our methodology aims to forecast the magnitude of an imminent earthquake based on the given input data, though it does not predict the timing of the event. This prediction model has undergone rigorous testing with selective historical and instrumental data to ensure its accuracy in forecasting both the magnitude and the location of the next earthquake, albeit without date and time specifics. For example, inputting data from the most recent recorded earthquake enables the model to predict the magnitude and site of a future earthquake. While predicting earthquakes without temporal information may appear to have limited immediate application, it marks a notable scientific advancement. These findings lay a solid foundation for future research endeavors focused on incorporating temporal elements into earthquake prediction models.

Table 6. Testing earthquakes for individual dates by LSTM and ANN algorithms.

Date	Latitude	Longitude	Depth	Target	LSTM	ANN
1995	39.4	71.0	5.0	5.1: 14	4.45:14	5.2:-
2006	39.53	73.65	5.0	5.0:17	4.6:17	5.0:-
2008	39.5	73.7	30.0	4.5:17	4.6:17	4.8:-
2011	39.55	73.1	30.0	4.8:16	4.6:16	4.6;-
2014	39.38	71.75	13.0	4.7:15	4.5:15	4.6:-

The architecture's training involved using a database comprising data up to the year 1994, inclusive of the corresponding 29 features, to predict seismic activity for 1995. This methodology was similarly applied to verify predictions for the years 2006, 2008, 2011, and 2014. The actual database data and zone numbers are presented in the last column. Training on these architectures resulted in a root mean square error (RMSE) of 0.12 for the ANN model and 0.14 for the LSTM model. While forecasting future earthquakes without temporal information may seem of limited practical use, it represents a significant scientific achievement. This result lays the groundwork for future research aimed at incorporating time-specific predictions.

As a complete seismological database for 2023 was not available, the study focused on monitoring the year 2022. For this, the LSTM model was trained separately for each zone, incorporating its unique characteristics and all recorded earthquakes within that zone of Ferghana Depression. This approach enabled the prediction of potential earthquakes for 2022, the details of which are documented in **Table 7**.

Table 7. Testing by LSTM algorithm earthquakes for 2022 year implementing a seismological database of Ferghana Depression

Forecast for 2022	
Zones	**Magnitude:Predicted/Real**
3	3.8/4.2
4	4.3/4.6
8	4.0/4.4
8	3.8/5.2
9	3.7/4.0
9	3.9/4.0
10	4.2/4.0
16	4.1/4.1
16	4.2/4.2

A review of the data tables reveals that deep learning successfully predicted the majority of earthquakes that occurred in 2022. The limitation of these deep learning algorithms in forecasting strong earthquakes with magnitudes greater than M>4.3, as shown in **Table 7**, can be attributed to the scarcity of high-magnitude events in the training dataset. A potential solution is to enrich the seismological database with more significant earthquakes from a broader region, encompassing the area under study, and subsequently retrain the models for more localized predictions. Implementing a seismological database that includes Central Asia and the Fergana depression enabled the prediction of earthquakes with magnitudes up to M=5.3, as indicated in **Table 8**.

Table 8. Predicted and real magnitudes and zones by LSTM algorithm for 2022 year implementing a seismological database of Central Asia

Year	Month	Day	Lat	Lon	Depth	Real Mag	Predict Mag	Real Zone	Predict Zone
2022	1	23	39,8	73,08	29	4,1	*4,9*		*10*
2022	3	14	40,3	70,18	27	*4,6*	5,2	*4*	8
2022	3	21	41,8	70,87	18	4,5	5,3	8	10
2022	3	31	41,5	73,06	10	5,2	4,1	3	7
2022	6	11	40,8	69,64	20	4,2	5,2	8	9
2022	6	23	41,4	72,78	11	4,7	4,1	16	9
2022	6	26	39,4	72,75	13	4,2	4,3	7	11
2022	7	27	41,8	71,74	5	4,4	4,9	3	9
2022	9	8	39,8	69,75	8	5,0	4,4	16	6
2022	11	1	39,9	72,94	10	4,4	*4*	16	*10*
2022	11	21	39,6	73,12	*8*	*4,4*	4,4	*10*	

Table 8 should be interpreted as follows: The machine was trained using data up to and including 2021, enabling verification of its predictions against the 2022 database. After the initial January 2022 event, a prediction request was submitted. The machine forecasted an earthquake with a magnitude of M=4.9 in zone No. 10. In reality, an earthquake occurred in March in adjacent zone No. 4, with a magnitude of M=4.6. These magnitudes and zones are highlighted in red in the table. Following this event, the machine predicted an earthquake in zone No. 8 with a magnitude of M=5.2, and so on. Overall, the machine accurately predicted 7 out of 11 events with a magnitude accuracy of M=0.3, except for one event noted as a second-order occurrence in Table 8. The slight discrepancy in exact zone prediction, despite the proximity of the zones, can be attributed to the initial division into 17 zones not being sufficiently precise. This highlights the need for selecting features based on morphologically homogeneous zones.

Nevertheless, our experience demonstrates that employing the methodologies described herein enables effective monitoring of strong earthquakes, particularly when the database is systematically updated with records of new seismic events, including weaker earthquakes. The logical progression for future work involves broadening the geographic scope of earthquake forecasting.

4. Conclusions

1. **Earthquake Prediction in the Fergana Depression:** Utilizing Cora-3, Cora-4, and Random Forest machine learning algorithms, we successfully identified potential locations for strong earthquakes in the Fergana depression.

2. **Identification of Geodynamic Features:** The study pinpointed the most relevant geodynamic features associated with strong earthquakes in the region.

3. **Forecasting Major Earthquakes:** We determined the possibility of future earthquakes with magnitudes M≥6.5 in specific zones of the Fergana depression.

4. **Prediction of Smaller Magnitude Earthquakes:** Earthquakes with magnitudes M≤4.3 in the Fergana depression were effectively predicted using LSTM and ANN machine learning algorithms.

5. **Comprehensive Earthquake Monitoring:** The research included detailed monitoring of earthquake activities throughout 2022 in the Fergana depression.

Author Contributions

- **Conceptualization and Methodology:** Initiated and designed the research framework — I.A., A.A., and R.L.

- **Software Development:** Contributed to the development and implementation of software tools — A.B. and J.M.

- **Writing, Review, and Editing:** Drafted the manuscript and performed critical revisions — I.A. and A.A.

- **Final Approval:** All authors have read and consented to the final version of the manuscript for publication.

Conflicts of Interest

The authors declare no conflicts of interest in relation to this article.

Acknowledgments

Conflict of interest

The authors declare no conflict of interest.

References

1. Ma L., Zhu L., Shi Ya. Attempts at using seismicity indicators for the prediction of large earthquakes by Genetic Algorithm-Neural Network method (1998), https://www.researchgate.net/publication/265293953
2. Ashif P., Hojjat A. Neural network model for earthquake magnitude prediction using multiple seismicity indicator//International Journal System. (2007), Vol. 17, No.1, Pp. 13-33.
3. Panakkat A., Adeli H. Recurrent neural network for approximate earthquake time and location prediction using multiple seismicity indicators.//Computer-Aided Civil and Infrastructure Engineering (2009), 24 (4), Pp. 280– 292.
4. Ashit K.D. Earthquake prediction using artificial neural networks//International Journal of Research and Reviews in Computer Science (IJRRCS) (2011), Vol. 2, No. 6, Pp. 2079-2557.
5. Mahmoudi J., Arjomand M. A., Rezaei M., Mohammadi M. H. Predicting the earthquake magnitude using the multilayer perceptron neural network with two hidden layers.//Civil Engineering Journal (2016), 2(1), Pp.1–12.
6. Asim K. M., Martínez-Álvarez F., Basit A., Iqbal T. Earthquake magnitude prediction in Hindukush region using machine learning techniques//Nat. Hazards (2017), Vol. 85, No. 1, pp. 471–486, doi: 10.1007/s11069-016-2579-3.
7. Rouet-Leduc, B., Hulbert, C.,Lubbers, N., Barros, K.,Humphreys, C. J., & Johnson, P. A. Machine learning predicts laboratory earthquakes.//Geophysical Research Letters (2017), 44, 9276–9282, https://doi.org/10.1002/2017GL074677
8. DeVries, P.M.R., Viégas, F., Wattenberg, M. Meade B.J. Deep learning of aftershock patterns following large earthquakes. //Nature (2018), 560, 632–634, https://doi.org/10.1038/s41586-018-0438-y
9. Kuyuk H.S., Susumu O. Real Time Classification of Earthquake using Deep Learning//Procedia Computer Science (2018), 140, Pp. 298-305, DOI: 10.1016/j.proc.2018.10.316
10. Shcherbakov, R. , Zhuang J., Zöller G., Ogata Y. Forecasting the magnitude of the largest expected earthquake//Nature Communications (2019), 10:4051, https://doi.org/10.1038/s41467-019-11958-4
11. Vardaan K., Bhandarkar T., Satish N., Sridhar S., Sivakumar R., Ghosh S. Earthquake trend prediction using long short-term memory RNN//International Journal of Electrical and Computer Engineering (IJECE) (2019), Vol. 9, No. 2, April, pp. 1304~1312 ISSN: 2088-8708, DOI: 10.11591/ijece.v9i2.
12. Bangar D.,Gupta D.,Gaikwad S., Marekar B., Patil J. Earthquake prediction using machine Learning Algorithms//Int. J. of Recent Technology and Engineering (2020), March, Vol.8, Issue 6, Pp. 4684-4688, DOI:10.35940/ijrte.E9110.018620

13. Salam M., Ibrahim L., Abdelminaam D. Earthquake Prediction using Hybrid Machine Learning Techniques//International Journal of Advanced Computer Science and Applications (2021), Vol. 12, No. 5, DOI: 10.14569/IJACSA.2021.0120578

14. Okazaki, T., Ito, T., Hirahara, K. Useda N. Physics-informed deep learning approach for modeling crustal deformation. Nat Commun (2022), 13, 7092, https://doi.org/10.1038/s41467-022-34922-1

15. Wang Q., Guo Y., Yu L., Li P. Earthquake Prediction based on Spatio-Temporal Data Mining: An LSTM Network Approach//IEEE, Transactions on Emerging Topics in Computing (2020), DOI:10.1109/TETC.2017.2699169

16. Gelfand I.M., Guberman Sh.A., Izvekova M.L., Keilis-Borok V.I., Ranzman E.Ya. On criteria for high seismicity//Dokl. Academy of Sciences of the USSR (1972), Vol. 202, No. 6, Pp. 1317–1320. [in Russion]

17. Bongard M. M., Vaintzvaig M. N., Guberman Sh. A., Izvekova M. L., Smirnov M. S. Using a learning program to identify oil-bearing reservoirs.//Geology and geophysics (1966), No. 6, Pp.96-105[in Russion]

18. Gvishiani A. D., Soloviev A. A., Dzeboev B. A. Problem of Recognition of Strong-Earthquake-Prone Areas: a State-of-the-Art Review//Izvestiya, Physics of the Solid Earth. (2020), Vol. 56, No. 1, Pp. 1-23, DOI:10.1134/S1069351320010048

19. Kossobokov V. G., Shchepalina P. D. Times of Increased Probabilities for Occurrence of World's Largest Earthquakes: 30 Years Hypothesis Testing in Real Time//Izvestiya, Physics of the Solid Earth (2020), Vol. 56, No. 1, Pp. 36-44. DOI: 10.1134/S1069351320010061

20. Shebalin P. N. The Rise of Earthquake Correlation Range and the Chains of Earthquakes before Large Seismic Events//Izvestiya, Physics of the Solid Earth. (2020), Vol. 56, No.1, Pp. 30-42, DOI:10.1134/S1069351320010139

21. Gzovsky M.V. Fundamentals of tectonophysics, M.: Nauka(1964), 536 p., [in Russian]

22. Riznichenko Yu.V. Problems of seismology, M.: Nauka(1985), 405 p. [in Russian]

23. Atabekov I.U., Sadykov Yu.M. Stress State of the Earth's Crust in the Western Tien Shan in Central Asia (Uzbekistan): A Mathematical Stress Model//Geotectonics (2022), No.3, pp. 306-320, DOI: 10.1134/S0016852122030037

24. Atabekov I. Earth Crust's Stresses Variation in Central Asian Earthquake's Region.//J. Geodesy and Geodynamics (2020),. Vol. 11, Issue 4, July, Pp. 293-299 https://doi.org/10.1016/j.geog.2019.12.005

25. Atabekov I. Numerical models of earthquake's mechanism//J. Geodesy and Geodynamics (2021), 12, Pp. 148-154, https://doi.org/10.1016/j.geog.2021.03.002

26. Atabekov I.U. Tectonic Earthquake Source Mechanism Model Based on Moment Theory//Izvestiya, Physics of the Solid Earth (2023), vol. 59, №1, Pp. 1–14, https://doi.org/10.1134/S1069351323010019

27. Haykin S.O. Neural Networks and Learning Machines. Third Edition, Pearson Education (2011), 936 p.