ANNALS OF GEOPHYSICS, 66, 6, SE636, 2023: doi: 10.4401/ag-8946
OPEN ACCESS

Earthquake forecast by imbalance machine learning using geophysical predictors

Tengiz Kiria, Tamaz Chelidze, George Melikadze, Tamar Jimsheladze, Gennady Kobzev

M. Nodia Institute of Geophysics, Tbilisi State University, Tbilisi, Georgia

Article history: received February 1, 2023; accepted October 10, 2023

Abstract

In the present paper we consider the earthquake forecast as a binary problem of machine learning on the imbalanced data base applied to five regions of Georgia. For the training we used geophysical data base collected in 2017-2021, namely, variations of statistical characteristics of geomagnetic field components, seismic activity, water level in deep boreholes and tides. In this version a new predictor – the weighted seismic activity for previous 5 days - $W(t)$ – is added compared to the predictors' list used in previous papers. Besides, the length of the used database is increased 3 times compared to the earlier results. As in the database the earthquakes of M > 3.5 are rare, the number of negative cases is large (there are many days without EQs of M > 3.5), meaning that there is a strong imbalance between positive and negative cases of the order of 1:20; we apply the specific methodology Matthews' correlation coefficient (MCC) and F1 score to avoid the strong imbalance effect.

Keywords: Earthquake forecast; Water level in wells; Geomagnetic variations; Micro-seismicity; Machine learning on imbalanced data; Receiver operating characteristics

1. Introduction

The seismic process is without doubt a complex process: according to the accepted definition, complexity appears in systems, which are composed of many components interacting nonlinearly. The complexity theory (nonlinear dynamics) approach, requires the detail knowledge of the real process, which allow to describe it by the system of differential equations. Complexity analysis allow revealing the existence of long-term correlations in the temporal, spatial and energy distributions in dynamical systems such as seismicity using mathematical models of the process [Chelidze et al., 2018]. On the other hand, last years appear modern machine learning (ML) approach, which is concerned with developing algorithms. ML methods improve their performance with increasing the volume of input information [Li, 2020]. This approach gained increasing attention in solving the problems, where it is impossible to formulate exact mathematical models but on the other hand there are a lot of real data measurements. The ML allow to create data-driven approach to understanding and forecast of behavior of many complex system without constructing the exact analytical model – in contrast to the complexity theory.

Last years, ML approach give a lot of promising results in forecasting both laboratory and natural earthquakes [Rouet-Leduc et al., 2017; Rouet-Leduc et al., 2018; Ren et al., 2020; Johnson et al., 2021]. It is interesting to note that one of the first publications devoted to application of ML for the EQ forecast belong to Chelidze et al. [1995]. In the paper authors applied the method of Generalized Portrait (now Support Vector Machine, SVM) suggested by Vapnik and Chervonenkis [1974], Vapnik [1984] to forecast the Caucasian EQs of magnitude 5 and more.

In the earlier paper [Chelidze et al, 2020] the problem of EQ forecast in the Caucasus region was considered using one-year only WL and geomagnetic observations. In the present paper a new predictor – the weighted seismic activity for the previous 5 days $W(t)$ – is added compared to the predictors used in the previous paper. The parameter $W(t)$ reflects the previous seismic activity in the chosen area for a given time interval and is used for improving the model, predicting future relatively strong ($M > 3.5$) events. Besides, in the previous paper, due to scarcity of data, we apply ML to any of chosen four regions and used fifth regions' data for testing. In the present paper, the length of the tested database is increased 3 times compared to earlier results; in this case the data for training/testing were divided in the ratio 70/30 for each station and learning was carried out for each of them independently.

In our region the ratio of time intervals containing EQs of $M > 3.5$ to the duration of seismically quiet periods is approximately 1:20, which mean that during a year the seismically active periods for EQs of $M > 3.5$ is approximately 20 times shorter than that of aseismic periods. Consequently, in the learning sequence there is a clear imbalance between seismo-active and quiet periods, which should be taken into consideration.

2. The network of observations, preparation of data bases and methodology of analysis

2.1 The network

M. Nodia Institute of Geophysics at Tbilisi State University serves the geophysical observation network including: geomagnetic variation, water level in deep wells, earth tides and temperature. The seismic network is served in recent years by the Seismological Monitoring Center of Ilia State University (Tbilisi). Analysis of the data show that there are certain connections between anomalies in the observed fields and seismic events of $M > 3.5$, which allow to develop an EQ forecasting ML model.

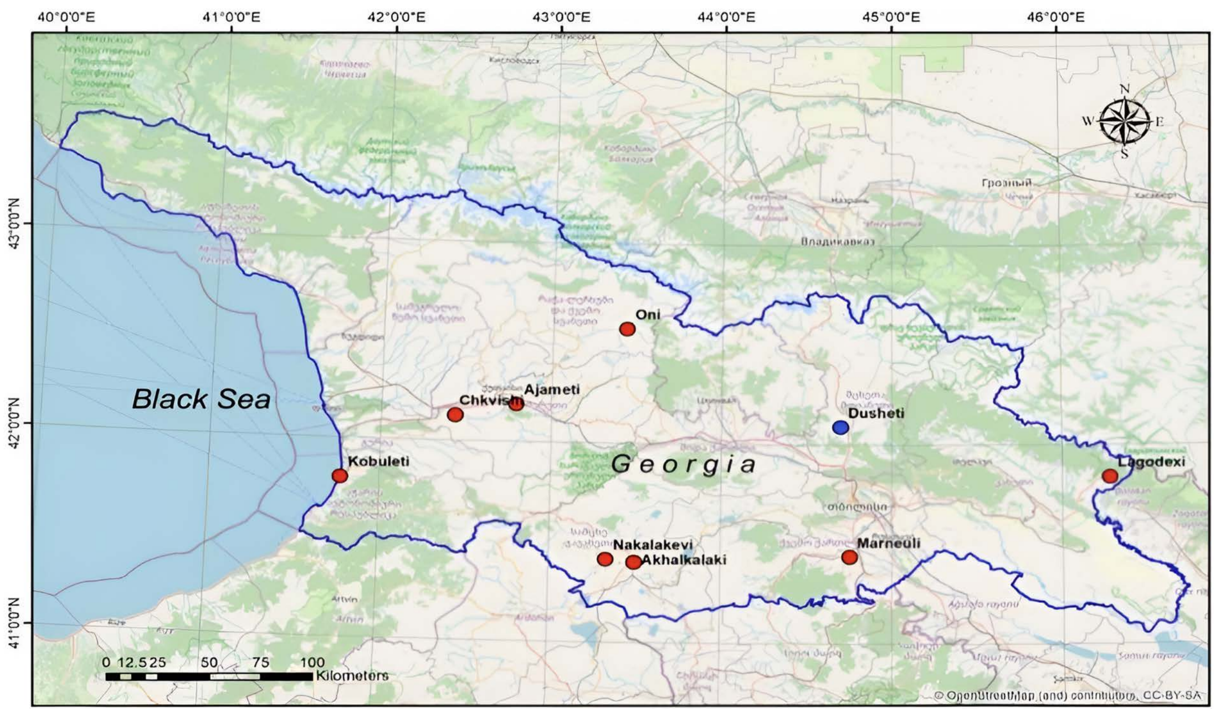

Figure 1. Map of Georgia with location of deep wells' network (red circles) for water level monitoring and Dusheti Geophysical Observatory.

WL monitoring network in Georgia include several deep wells, drilled in confined sub-artesian aquifers: data obtained at stations Kobuleti, Akhalkalaki, Ajameti, Chakvi, Lagodekhi were used in this paper (Figure 1). Measurements are done by sensors MPX5010 with resolution 1% of the scale (company Freescale Semiconductors; www.freescale.com) and recorded by data logger XR5 SE-M (company Pace Scientific; http://www.pace-sci.com/data-loggers-xr5.htm). The sampling rate at all these wells was 1/min. The data are transmitted remotely by the modem Siemens MC-35i Terminal using program LogXR. Variations of water level represent an integrated response of the aquifer to different periodic and quasi-periodic (tidal variation, atmosphere pressure) as well as to non-periodic influences – earthquakes. EQ-related strains in the earth crust are of the order of 0.1-0.001 μm. The atmosphere pressure factor was subtracted from the summary WL variations.

Magnetic data (x, y, z components) were obtained at Dusheti Geophysical Observatory (Lat 42.052N, Lon 44.42E), with the fluxgate magnetometer FGE-95 (Japan); the count rate was 1/sec and accuracy 0.1 nT.

Anomalies in hydraulic [Wang and Manga, 2010] and geomagnetic data [Balasis et al., 2011; Zotov et al., 2013; Buchachenko, 2021] as the phenomena observed in the precursor manifestation area R. Here we choose the radius of manifestation area (interaction length) R = 200 km for hydraulic precursors of EQs of M > 3.5 for a given well. There are different assessments of EQ precursors' area [Dobrovolsky et al., 1979, Pregean and Hill, 2009]. We presume that there are at least two physical mechanisms, which can explain the accepted long radius of action of hydraulic precursors (R = 200 km): i. poroelastic effect and ii. fast squirt-flow of pore water from the future EQ source to the well [Dvorkin et al., 1994]. Such signals can travel a long distances [Chelidze et al., 2019]. We accept that the interaction length R for geomagnetic precursors of the order of 300 km around Dusheti observatory.

2.2 Preparation of databases and methodology of analysis

We consider our work as a ML problem aimed at solving a binary classification task of detecting regions/periods, where the M > 3.5 EQ probability is high or low. Analysis of our regular geophysical observations (1 count per min) and 5-days seismicity characteristics lead to conclusion on existence of anomalies in statistical assessments of these data before M > 3.5 EQs. Namely, some days before EQs the above geophysical parameters became unstable and form time sequences with high dynamical characteristics. In order to reveal the abovementioned connection, we carried out pre-formatting of the observed data for application of machine learning tools, where we take the data of the previous five days as the known statistics (Input) and as the Output – the occurrence or absence of the M > 3.5 event as the sixth day after every five days. In other words, we train the program during 5 days in order to predict the 6-th days' seismicity, namely, occurrence of M > 3.5 EQ or absence of such event.

Now we define the way, in which the new seismicity predictor is calculated. For each time moment t_i the following expression, assessing the local seismicity is calculated for the previous 5-days window:

$$W(t_i) = \frac{\sum_{i-5*1440<j<i}\left[M_j \cdot (t_i - t_j)\right]}{\sum_{i-5*1440<j<i}(t_i - t_j)} \tag{1}$$

here M_j is the magnitude of a given EQ in the catalog for the moment t_j, $t_i - t_j$ is the time interval before this EQ; i and j have the 1-minute discretization; 5*1440 is the 5-days duration running window expressed in minutes; the i value varied from 5*1440 to 2,102,400 for a given region. As a result of calculations by (1) we obtain the value of $W(t)$ for all t_i moments. From the equation (1) it is evident that the longer is the time interval in the 5-day window without EQ, the less is its weight in the value of $W(t_i)$. Quite often $W(t_i) = 0$, as during 5-days period there does not occur any EQ. So, actually, the $W(t_i)$ depend on the short-term foreshock activity before M > 3.5 EQ; the short-term activation of weak seismicity before strong EQ is one of accepted predictors [Artikov et al., 2018; Saccorotti, 2022].

According to (1), the value $W(t)$ is sometimes a large number, so the expression 1-1/$W(t)$ converts it to the range [0:1], which is more convenient for graphical representation of the results.

In Figure 2 we present the probabilistic-statistical dependence between normalized values of function 1-1/$W(t)$ and EQ occurrence. In other words, if the occurrence of EQ events is correlated in time and space (namely, if in the 5-days window one or more EQ of magnitude 3.5 is fixed), the probability of EQ in the nearest future increases.

Thus, it seems that calculation of function $W(t)$ and its inclusion in the EQ forecast model should improve the results.

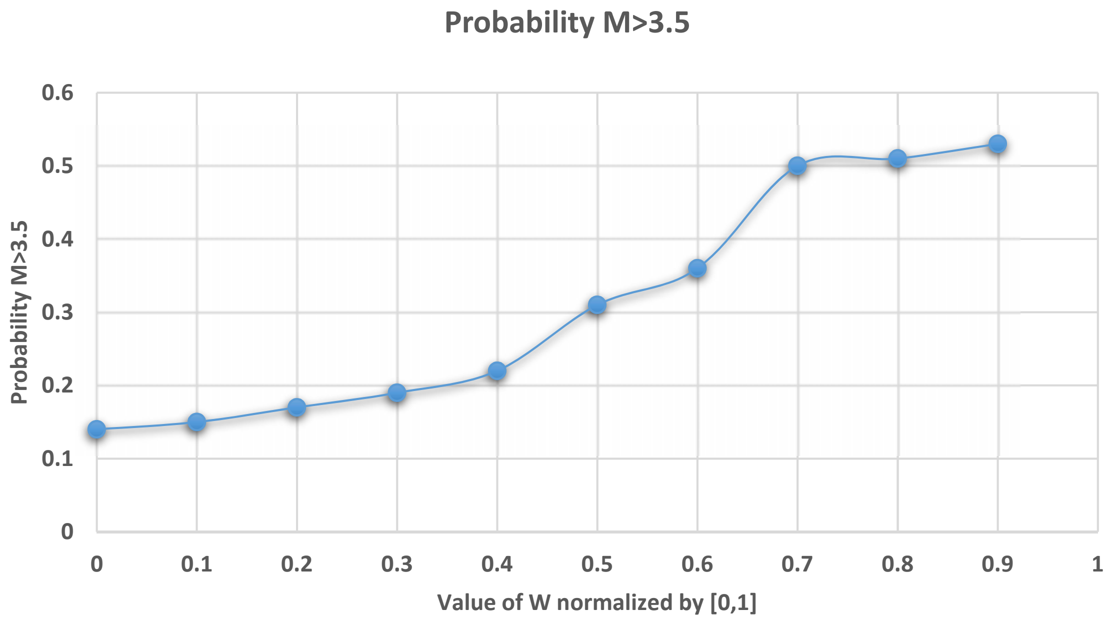

Figure 2. The dependence between 5-days values of $W(t)$ and probability of frequency of EQ M > 3.5 occurrence obtained for the next 6-th day (i.e. the value of $W(t)$ normalized in the interval [0,1] in order to avoid large values of $W(t)$).

2.3 Structure of the learning base and machine learning application principles

Algorithms of ML are often used to forecast complex processes, especially when data are very big. In our previous papers, the analysis was founded only on the one-year WL and geomagnetic data collected for five regions of Georgia [Raschka, Mirjalili, 2019]. In the present paper the volume of input data increases 5-folds; besides, the preliminary seismicity parameter $W(t)$ is added to predictors' list.

The systematic predictive signs were revealed in the spectral and different statistical parameters of observed fields' time series. Below we present the scheme illustrating the whole cycle of training/testing operations for performing ML-based forecast (Figure 3).

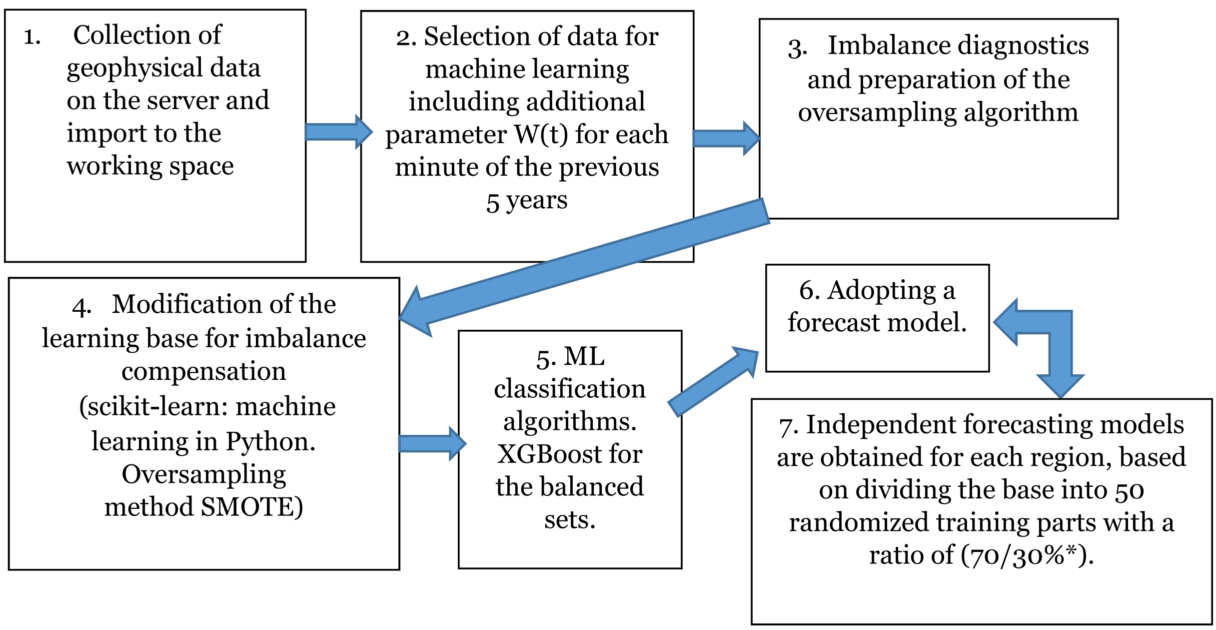

Figure 3. The block-scheme of actions in the machine learning EQ forecast method.

In the accepted ML model, we divide experimental data at the ratio 70/30%, where 70% of the data are used for training and remaining 30% – for testing. This approach is widely used to create a set of independent forecasting models for different regions [Raschka and Mirjalili, 2019]. The use of multiple 70/30% splits helps to assess the robustness of the models and their ability to generalize different data scenarios. This operation serves to give statistical significance to above procedure – the randomization of training parts helps in creating a diverse set of training and testing data. This is important to ensure that each split is statistically significant and representative of the region's data distribution. Besides, we used hyperparameter tuning: since we use 50 training models per region, the operation become crucial to achieve optimal model performance. Each model might perform differently based on factors like learning rate, regularization strength, etc. So the code for the XGBoost is:

XGBoost Specify hyperparameters = {'objective': 'binary:logistic', 'max_depth': 3, 'learning_rate': 0.1, 'n_estimators': 100, 'subsample': 0.7, 'colsample_bytree': 0.8, 'gamma': 0.1, 'alpha': 0.1, - L1 regularization, 'lambda': 1 - L2 regularization}.

In Table 1, we present the input data for ML as well as the output EQ data: in total, there are 29 inputs and one output – for data fields nomination. The mentioned inputs present statistical values of WL, seismic and geomagnetic predictors in 5-day windows. Since we have every minute observations (1440 data set per day), we create the following table for different statistical parameters – average, Max, Min, Std Dev of these data as well as for the seismic criterion $W(t)$, determined by formula (1). The terms in the Table 1 are shortened comments on the previous 5 days' values of the parameters used – for example: average of MagnitX is the average value of the X component of magnetic field for the previous five days, Max of water – is the maximum value of water level for the previous five days, etc. Note that the weight of the suggested seismicity qualifier is quite high (Figure 4).

Inp_1	Average of MagnitX	Inp_16	Min of MagnitY
Inp_2	Average of MagnitY	Inp_17	Min of MagnitZ
Inp_3	Average of MagnitZ	Inp_18	Min of Tid1
Inp_4	Average of Tid1	Inp_19	Min of Tid2
Inp_5	Average of Tid2	Inp_20	Min of Tid3
Inp_6	Average of Tid3	Inp_21	Min of water
Inp_7	Average of water	Inp_22	StdDev of MagnitX
Inp_8	Max of MagnitX	Inp_23	StdDev of MagnitY
Inp_9	Max of MagnitY	Inp_24	StdDev of MagnitZ
Inp_10	Max of MagnitZ	Inp_25	StdDev of Tid1
Inp_11	Max of Tid1	Inp_26	StdDev of Tid2
Inp_12	Max of Tid2	Inp_27	StdDev of Tid3
Inp_13	Max of Tid3	Inp_28	StdDev of water
Inp_14	Max of water	Inp_29	$W(t)$
Inp_15	Min of MagnitX		

Table 1. The structure of learning database.

We used the XGBoost algorithm for machine learning; this allow calculate importance of features using the built-in methods provided by the XGBoost library [Raschka, Mirjalili, 2019]. The plot of importance function provided by XGBoost creates a bar chart in Figure 4 showing the importance scores of features based on their contribution to the model's performance. The importance scores are calculated using the "weight" metric by default, which represents the number of times a feature is used to split the data across all trees in the ensemble. We can customize the metric using the importance_type parameter of the plot_importance function.

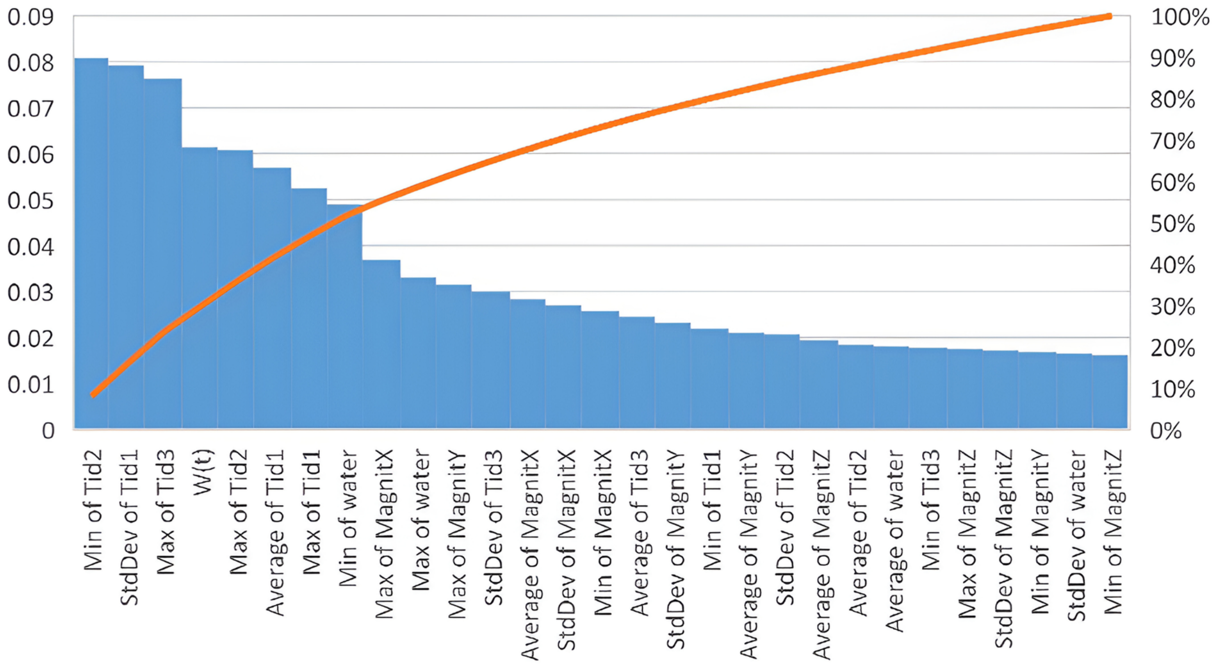

Figure 4. Weights of different statistical measures obtained in the process of machine learning using average input data for all five regions (blue columns). The red curve shows the cumulative percent of separate weights. Note that the weight of the suggested seismicity qualifier is quite high.

3. Machine Learning algorithms

3.1 Solution of imbalance problem and algorithm training

The ML algorithm was initially designed for the analysis of the balanced datasets, where, in the bivariate case, the number of 0 and 1 events is almost equal. In reality, in many problems, the data sets are imbalanced and without application of special methods, the decision of machine will be biased. Namely, the minority class, which can be the most interesting one in such problems as medicine, business, catastrophe forecasts can be ignored at all [Fernandez et al., 2018; Mena and Gonzales, 2006; Johnson and Khoshgoftaar, 2019; Malik and Ozturk, 2020; Brunton and Kutz, 2022].

As we mentioned earlier, in Georgia the ratio of time intervals containing EQs of $M > 3.5$ to “quiet” interval (time intervals without EQs of $M > 3.5$) is 1:15-1:20, i.e. there is strong enough imbalance between minority – “seismo-active” and majority – “quiet” cases. It is known that in order to overcome the imbalance we have to compensate its effect [Chawla et al, 2018; Brownlee, 2021], which can be avoided by special technique – well planned artificial multiplication of strong seismic events in the learning sequence. This procedure is called oversampling in the ML approach [He and Garcia 2009, Brownlee, 2021]. In the paper, we present the results of ML approach applied to learning sequences composed by the above technique.

It is known that chaotic impact on the training data lead to grave mistakes in ML, especially when one deals with the necessity of data oversampling. We have run into the problem due to strong imbalance in the significant EQs’ data. It is important to note that correction of imbalance should be done before learning procedure only on the randomly chosen part of the whole training set and only 30% of data should be left as the testing part.

We used special library of Python – imblearn.over_sampling SMOTE for learning. The results show that despite complexity of data randomization of training samples did not affect significantly MCC results. As the learning algorithm the method: XGBoost for the Balanced (Oversampled) Sets was used.

3.2 Analysis of the obtained results; Choosing appropriate metrics

After some tests it became evident that in order to find some useful joint "memory" in different time series we need to use the deep learning algorithms (DLA). We apply the stochastic optimization scheme ADAM (Kingma and Ba, 2014) for a joint data analysis: this approach leads to optimization of the target function by stochastic gradient approach using combination of algorithms [Karpathy, 2017].

In order to correctly assess the effectiveness of ML algorithms in the case of strong imbalance between majority and minority classes it is recommended to use Matthews correlation coefficient MCC or F1 score measure [Matthews, 1975; Chicco and Jurman, 2020; Chicco et al., 2021; Chelidze et al., 2022]:

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \tag{2}$$

where *TP*, *TN*, *FP*, *FN* are correspondingly correctly predicted positives, correctly predicted negatives, wrongly predicted positives and wrongly predicted negatives. The MCC coefficient varies in the range (−1, 1); the model is optimal if the MCC = 1. MCC is a measure, which can be used even for imbalances data, i.e. when the numbers of negative and positive classes are very different. MCC close to (+1) provide high values of all main parameters of confusion matrix.

At the same time, according to Chicco et al., [2021], if one of the classes (say, positive ones) is more important than negative cases, the F1 can be the preferable score. The F1 score is a harmonic mean of Precision and Recall. Precision and Recall are preferable when the data are imbalanced as they take into account both types of errors (false negatives as well as false positives). The resulting single metric works well on imbalanced data:

$$F1 = 2TP/(2TP + FP + FN) = 2 \times precision.recall/(precision + recall) \tag{3}$$

$F1$ varies in the range (0, 1). The maximum value 1 is reached for the perfect positive classification, i.e. when $FN = FP = 0$ and the minimum value 0 is attained at $TP = 0$.

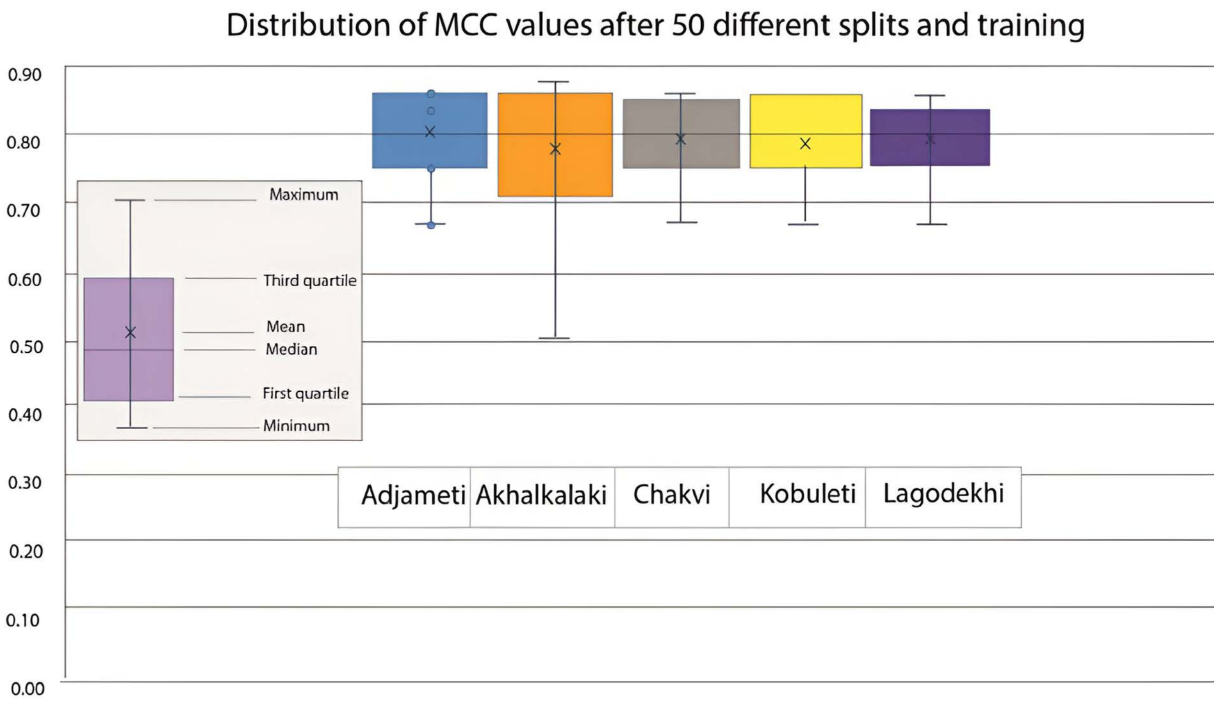

Figure 5. The range of variation of ML results for Matthews' correlation coefficient for 5 regions after 50 randomization tests for 70/30 divided training/testing sequences. The comments on meaning of different marks in the graph are presented in the window on the left.

As mentioned above, we prepared independent learning-training tables for 5 regions and carried out the learning procedure for each of them separately. The table length for each region was 5*365 = 1825; 70% of these data was used for training, and 30% – for testing procedure. After this, the data for each region were randomized 50 times and training and learning procedure were carried out on the randomized data sets. Figure 5 shows the results of learning on the randomized sequences using MCC. In the specific, figure illustrates the range where the Matthews correlation coefficient varies due to the randomization. It is evident that the width of distribution of MCC values differs for various regions, but generally, even after many randomizations tests the results are quite stable, which seems encouraging.

It's crucial to note that we do not balance the test data. Balancing test data would be counterproductive, as it might lead to misleading results. We partitioned the training data 50 times, balanced the training set each time, and then made predictions on this test set. The outcomes were consistent with only slight variations (Figure 5). Note, when we organize the data based on the study period – i.e. training on the data from January to September and testing from October to December – there are no significant differences in the results.

It is evident that increasing data length (time interval) does not change our preliminary estimates. Using the results of the above analysis, we compiled the confusion matrix (Table 2) forecasting EQs of $M \geq 3.5$ around the Axalkalaki, Adjamerti, Chakvi, Kobuleti and Lagodekhi boreholes for the years 2017-2021.

Adjameti			
TN	**428**	**34**	FN
FP	**27**	**58**	TP
Akhalkalaki			
N	**414**	**49**	FN
FP	**19**	**66**	TP
Chakvi			
TN	**420**	**44**	FN
FP	**24**	**60**	TP
Kobuleti			
TN	**435**	**45**	FN
FP	**19**	**49**	TP
Lagodekhi			
TN	**451**	**31**	FN
FP	**22**	**44**	TP

Table 2. The ML confusion matrix results for 30% randomized data from five regions.

The forecast results were assessed using MCC, F1 score and Accuracy values – Table 3, where we present ML results for different statistical measures obtained at 5 stations.

Table 3 shows that the results for F1 score, which is the combination of Precision and Recall, lead to estimates 0.85 ± 0.010, close to MCC values – 0.8 ± 0.012.

As was expected, the Accuracy assessment seems to be too optimistic. This is the result of strong imbalance in the input data – namely, to large values of TN [Chelidze et al., 2022]. As a result, the minority class (strong EQs) is practically ignored in Accuracy assessment.

	Matthews Coef.	Accuracy	Precision	Recall	F1 score
Adjameti	0.812	0.934	0.890	0.820	0.853
Akhalkalaki	0.788	0.921	0.873	0.809	0.840
Chakvi	0.798	0.927	0.863	0.828	0.845
Kobuleti	0.804	0.925	0.903	0.807	0.853
Lagodekhi	0.798	0.926	0.880	0.813	0.845

Table 3. The averaged results of ML for different regions.

4. Testing MCC results on randomized EQ catalog

For testing the validity of MCC and F1 score approach, we randomized the input training data. Namely, in the majority (aseismic) data sets we implanted randomly the additional seismic events according to the following rule: in the middle of the quiet-days cluster, we place one event of $M > 3.5$. According to the theory, MCC for randomized data sets should decrease significantly. Below we show the results of MCC testing on the randomized 5 regions' data. The measures were taken to avoid the imbalance effect and exclude the overfitting possibility. After regularization (balancing) of the training data, we build the Confusion Matrix and performed Receiver Operating Classification in order to forecast the next day probability of $M > 3.5$ earthquake occurrence. We found that after randomization of EQ dates in the training dataset both MCC and F1 score decrease significantly, when Accuracy stays too optimistic (Table 4).

	Matthews Coef.	Accuracy	Precision	Recall	F1 score
Adjameti	0.564	0.888	0.658	0.602	0.629
Akhalkalaki	0.594	0.873	0.773	0.576	0.660
Chakvi	0.568	0.874	0.713	0.579	0.639
Kobuleti	0.548	0.883	0.721	0.521	0.605
Lagodekhi	0.576	0.903	0.672	0.592	0.629

Table 4. Training results for all five regions after earthquake events' time randomization.

5. Conclusions

In the present paper the EQ forecast problem for the West Caucasus region is considered, using data on geomagnetic variation, water level in deep wells, earth tides as well as additional predictive seismological parameter $W(t)$ on 3-years long learning/testing period. $W(t)$ characterize the EQ activity in the 5-days intervals before events of $M > 3.5$.

Besides, special attention was paid to compensate the imbalance effect leading to overfitting of data. We show that by application of the oversampling approach it is possible to obtain balanced assessments.

The confusion matrix was obtained, which show that such statistical measures, as Matthews correlation coefficient and F1 score give good results in forecasting regional events of $M > 3.5$, namely MCC in the range 0.8 ± 0.012 and F1 score in the range 0.85 ± 0.010. After randomization of the EQ catalog, the values of both MCC and F1 score decrease considerably.

References

Artikov, R.S., ... (2018). Identification of expected seismic activity areas by forecasting complex seismic-mode parameters in Uzbekistan. Geodesy and Geodyn., 2, Issue 2, Pages 121-130.

Balasis, G., I. A. Daglis, A. Anastasiadis, C. Papadimitriou, M. Mandea and K. Eftaxias (2011). Universality in Solar Flare, Magnetic Storm, and Earthquake Dynamics Using Tsallis Statistical Mechanics, Physica A, 390, 341-346, doi:10.1016/j.physa.2010.09.029

Brownlee, J. (2021). Imbalanced Classification with Python: Better Metrics, Balance Skewed Classes, Cost-Sensitive Learning. Kindle Edition. Amazon. Available at: https://amzn.to/3s4Ihfb

Brunton, S. and J. Kutz (2022). Data-Driven Science and Engineering, 2nd Ed. Cambridge UP.

Buchachenko, A. (2021). Magnetic Control of the Earthquakes. Open Journal of Earthquake Research, 10, 138-152, doi:10.4236/ojer.2021.104009

Chawla, N., F. Herrera, S. Garcia and A. Fernandez (2018). SMOTE for Learning from Imbalanced Data: Progress and Challenges, Marking the 15-year Anniversary, J. Artif. Intel. Res., 61, 863-905, doi:10.1613/jair.1.11192

Chelidze, T., G. Melikadze, T. Kiria, T. Jimsheladze and C. Kobzev (2020). Statistical and Non-linear Dynamics Methods of Earthquake Forecast: Application in the Caucasus, Front. Earth Sci., 8, 194, doi:10.3389/feart.2020.00194

Chelidze T., G. Melikadze, G. Kobzev, I. Shengelia, N. Jorjiashvili, E. Mepharidze (2019). Hydrodynamic and Seismic Response to Teleseismic Waves of Strong Remote Earthquakes in Caucasus, Acta Geophysica, 67, 1-16, doi:10.1007/s11600-018-00241-7

Chelidze, T., G. Sobolev, Y. Kolesnikov and A. Zavjalov (1995). Seismic Hazard and Earthquake Prediction Research in Georgia, J. Georgian Geophys. Soc., 1A, 7-39.

Chelidze T., F. Valliantos and Telesca (Editors) (2018). Complexity of Seismic Time Series. Amsterdam, Netherlands: Elsevier. https://shop.elsevier.com/books/complexity-of-seismic-time-series/chelidze/978-0-12-813138-1

Chelidze T., T. Kiria, G. Melikadze, T. Jimsheladze and G. Kobzev (2022). Earthquake Forecast as a Machine Learning Problem for Imbalanced Datasets: Example of Georgia, Caucasus, Front. Earth Sci., 10:847808. doi: 10.3389/feart.2022.847808

Chicco, D. and G. Jurman (2020). The Advantages of the Matthews Correlation Coefficient (MCC) over F1 Score and Accuracy in Binary Classification Evaluation, BMC Genomics, 21, 6, doi:10.1186/s12864-019-6413-7

Chicco, D., N. Tutsch and G. Jurman (2021). The Matthews Correlation Coefficient (MCC) Is More Reliable Than Balanced Accuracy, Bookmaker Informedness, and Markedness in Two-Class Confusion Matrix Evaluation, BioData Mining, 14, 13, doi:10.1186/s13040-021-00244-z

Dobrovolsky, I., S. Zubkov and V. Miachkin (1979). Estimation of the Size of Earthquake Preparation Zones, Pure Appl. Geophys., 117, 1025-1044, doi:10.1007/bf00876083

Dvorkin, J., R. Nolen-Hoeksema and A. Nur (1994). The Squirt-flow Mechanism: Macroscopic Description, Geophysics, 59, 428-438. doi:10.1190/1.1443605

Fernández, A., S. García and M. Galar (2018). Learning from Imbalanced Data Sets. Heidelberg, Germany: Springer, doi:10.1007/978-3-319-98074-4

He, H. and E.A. Garcia (2009). Learning from Imbalanced Data, IEEE Trans. Knowledge Data Engin., 21, 9, 1263-1284, doi:10.1109/tkde.2008.239

Johnson, J. and T. Khoshgoftaar (2019). Survey on Deep Learning with Class Imbalance, J. Big Data, 6, 27, doi:10.1186/s40537-019-0192-5

Johnson, P.A., B. Rouet-Leduc, L. Pyrak-Nolte, C.J. Marone, C. Hulbert, A. Howard, et al. (2021). Laboratory Earthquake Forecasting: A Machine Learning Competition, PNAS, 118, 5, e2011362118, doi:10.1073/pnas.2011362118

Karpathy, A. (2017). A Peek at Trends in Machine Learning. Available online at: https://medium.com/@karpathy/a-peek-at-trends-in-machine-learningab8a1085a106 (accessed June 12, 2020).

Kingma, D. and J. Ba (2014). Adam: A Method for Stochastic Optimization, Available online at: http://arxiv.org/abs/1412.6980http://arxiv.org/abs/1412.6980 (last accessed May 12, 2020).

Li, L., L. Li, O. Wanli, … (2020). Deep Learning for Generic Object Detection: A Survey, Int. J. Compu. Vision, https://doi.org/10.1007/s11263-019-01247-4

Malik, N. and U. Ozturk (2020). Rare Events in Complex Systems: Understanding and Prediction, Chaos, 30, 090401.

Matthews, B.W. (1975). Comparison of the Predicted and Observed Secondary Structure of T4 Phage Lysozyme, Biochim. Biophys. Acta (BBA) Protein Struct., 405, 2, 442-451, doi:10.1016/0005-2795(75)90109-9

Mena, L. and J. Gonzalez (2006). Machine Learning for Imbalanced Datasets: Application in Medical Diagnostic, in Proceedings of the Nineteenth International Florida Artificial Intelligence Research Society Conference, Melbourne Beach, Florida, USA, May 11-13, 2006. Available at: https://www.researchgate.net/publication/221438576

Prejean, S. and D. Hill (2009). Dynamic Triggering of Earthquakes, in Encyclopedia of Complexity and Systems Science, Editor A. Meyers (Berlin: Springer), 2600-2621. doi:10.1007/978-0-387-30440-3_157

Raschka, S. and V. Mirjalili (2019). Machine Learning and Deep Learning with Python, scikit-learn, and TensorFlow 2, Packt Publishing Limited, 770.

Ren, C.X., C. Hulbert, P.A. Johnson and B. Roulet-Leduc (2020). Machine Learning and Fault Rupture: A Review, in Advances in Geophysics, Amsterdam, Netherlands: Elsevier, 61. Available at: https://www.researchgate.net/publication/342216700

Rouet-Leduc, B., C. Hulbert, N. Lubbers, K.M. Barros, C. Humphreys and P.A. Johnson (2017). Machine Learning Predicts Laboratory Earthquakes, Geophys. Res. Lett., 44. doi:10.1002/2017GL074677

Rouet-Leduc, B., C. Hulbert, D.C. Bolton, C.X. Ren, J. Riviere, C. Marone, … (2018). Estimating Fault Friction From Seismic Signals in the Laboratory. Geophys. Res. Lett., 45, 3, 1321-1329. doi:10.1002/2017GL076708

Saccorotti G., R. Bruni and M. Bonini (2022). Recent Seismic Sequences and Activation of Normal Fault Systems in the Mugello Basin and Surrounding Areas (Northern Apennines, Italy), Front. Earth Sci., Solid Earth Geophysics, https://doi.org/10.3389/feart.2022.879160

Vapnik, V. and A. Chervonenkis (1974). Pattern Recognition Theory, Statistical Learning Problems, Moskva: Nauka Publisher.

Vapnik, V. (Editor) (1984). Algorithms and Programs for Recovering Dependences. Moscow: Nauka Publisher.

Wang, C.-Y. and M. Manga (2010). Earthquakes and Water. Heidelberg, Germany: Springer, 225.

Zotov, O., A. Guglielmi and A. Sobisevich (2013). On Magnetic Precursors of Earthquakes, Izvestiya, Phys. Solid Earth, 49 (6), 882-889. doi:10.1134/s1069351313050145

***CORRESPONDING AUTHOR: Tengiz KIRIA**,
M. Nodia Institute of Geophysics, Tbilisi State University, Tbilisi, Georgia
e-mail: kiria8@gmail.com