Machine Learning Methods for Earthquake Prediction: a Survey

Alyona Galkina
Saint Petersburg State University
Saint Petersburg, Russia
id.a.brickman@gmail.com

Natalia Grafeeva
Saint Petersburg State University
Saint Petersburg, Russia
n.grafeeva@spbu.ru

Abstract — **Earthquakes are one of the most dangerous natural disasters, primarily due to the fact that they often occur without an explicit warning, leaving no time to react. This fact makes the problem of earthquake prediction extremely important for the safety of humankind. Despite the continuing interest in this topic from the scientific community, there is no consensus as to whether it is possible to find the solution with sufficient accuracy. However, successful application of machine learning techniques to different fields of research indicates that it would be possible to use them to make more accurate short-term forecasts.**

This paper reviews recent publications where application of various machine learning based approaches to earthquake prediction was studied. The aim is to systematize the methods used and analyze the main trends in making predictions. We believe that this research will be useful and encouraging for both earthquake scientists and beginner researchers in this field.

Keywords — *earthquake prediction, data mining, time series, neural networks, seismology*

I. INTRODUCTION

At present, many processes and phenomena affecting different areas of human life have been studied enough to make predictions. Risk analysis makes it possible to determine whether the event is likely to occur at given period of time, as well as promptly respond to this event or even prevent it. However, even in the modern world there are events that we cannot influence. Such events, in particular, include natural disasters: tsunamis, tornadoes, floods, volcanic eruptions, etc. Human beings cannot stop the impending threat; but precautionary measures and rapid response are potentially able to minimize the economical and human losses.

However, not all natural disasters are equally well studied and "predictable". Earthquakes are one of the most dangerous and destructive catastrophes. Firstly, they often occur without explicit warning and therefore do not leave enough time for people to take measures. In addition, the situation is compounded by the fact that earthquakes often lead to other natural hazards such as tsunamis, snowslips and landslides. They may even cause industrial disasters (for instance, Fukushima Daiichi nuclear disaster was initiated by the Tōhoku earthquake that occurred near Honshu Island on 11 March 2011 and was the most powerful earthquake ever recorded in Japan [1]).

All these facts make the problem of earthquake prediction critical to human security. Since the end of XIX century, researchers in seismology and related branches of science have tried to discover so-called *precursors*, anomalous phenomena that occur before seismic events. Many possible precursors have been studied, including foreshocks (quakes which occur before larger seismic events), electromagnetic anomalies called "earthquake lights", changes of groundwater levels and even unusual animal behaviour. In some cases precursor appearance led to timely evacuation of civilians [2].

It is important to note that it is hard to use precursors for short-term forecasting, as they are they are not only characteristic of earthquakes (for instance, unusual lights in atmosphere may appear before geomagnetic storms or have a technogenic origin). Furthermore, different precursors preceded the quakes, which had different nature, occurred in different seismic zones and even seasons.

Thus, optimistic attitude towards the possibility of timely detection of earthquake hazards, which emerged in the 1970s because of a number of successful "predictions", have been replaced by skepticism [3]. This happened primarily because of numerous high-profile cases of wrong predictions [4]. Another reason was that no statistically significant precursors were found [5].

Currently there is no general methodology for earthquake prediction. Moreover, there is still no consensus in science community on whether it is possible to find a solution of this problem. However, rapid development of machine learning methods and successful application of these methods to various kinds of problems indicates that these technologies could help to extract hidden patterns and make accurate predictions.

These tendencies fully explain the amount of papers where the applicability of various machine learning algorithms to the the tasks of earthquake science is studied. Some of them are focused on precursor study: for instance, in paper [6] random forest algorithm is applied to acoustic time series data emitted from laboratory faults in order to estimate the time remaining before the next "artificial earthquake". Another application is discovering patterns of *aftershocks* which are small quakes that follow a large earthquake (referred to as a *mainshock*) and occur in the same area. One of the most recent examples is paper [7], where an artificial neural network in trained on more than 130.000 mainshock-aftershock pairs in order to model aftershock distribution and outperforms the classic approach to this task. However, although these fields of research are both very interesting and potentially helpful for solving the problem of earthquake prediction, the task formulated in the papers differs from the original one defined by seismologists (the definition is given in Section II), and therefore the results of these studies cannot be fully compared with the others.

However, despite the undoubted relevance of the problem, the whole time the research have been conducted, only a few authors have tried to systematize knowledge from various sources. In particular, one recent survey on a similar topic was found, published in CRORR Journal in 2016 [8]. The paper reviews using artificial neural networks for short-term earthquake forecasting. However, it is focused only on a single aspect of the problem: the authors mostly discussed various architectures and topologies of neural network models used to solve the problem. Therefore, the paper refers mainly to a limited group of specialists. The main objective of our review

is, on the contrary, to try to narrow the gap between seismology and computer science, as well as to encourage further research in this area. That is why this paper will attempt to cover all the main parts of a process of making predictions, including the search and preprocessing of earthquake data, the principles of feature extraction, as well as the methods of assessing the performance of machine-learning based predictors.

II. DESCRIPTION OF THE TASK

Despite words "forecast" and "prediction" are often used interchangeably, in earthquake science it is customary to distinguish them. Particularly, in [9] the idea was expressed that an earthquake prediction implies greater probability than an earthquake forecast; in other words, *a prediction is more definite than a forecast*, it requires greater accuracy. Therefore, it is worth noting that in this study we will deal mainly with earthquake prediction, since it seems to be more important from a practical point of view.

According to [10], the following information is required from the prediction of an earthquake in its simplest interpretation:

- a specific location;

- a specific time interval;

- a specific magnitude range.

Importantly, all of these parameters should be defined in such a way that *one could objectively state that some future earthquake does or does not satisfy the prediction*. It is necessary for both using and evaluating predictions. In particular, it is required to define "location" clearly and determine the exact spatial boundaries of the area, since an earthquake does not occur at a point.

Besides, the prediction is more useful and statistically verifiable if it includes the probability that the event that meets all above-mentioned criteria will occur [11]. That is, a prediction should specify *where*, *when*, *how big* the predicted earthquake is, and *how probable* is that it will occur in actual.

However, despite the importance of the problem of earthquake prediction and the existence of precise criteria that its solution should satisfy, there is still no general method for predicting earthquakes with sufficient accuracy. One of the main reasons is that it is extremely hard to build an accurate model of the process of earthquake occurrence. That is due to several reasons:

- Not all factors that may play roles in earthquake occurrence are discovered;

- Even well-known factors, such as the accumulated stress or seismic energy release rate, cannot be directly measured (or it is too hard to do it);

- The relationships between the occurrence of new earthquakes and these seismic features are shown to be complicated and highly non-linear.

All this leads to the use of increasingly complex methodologies when trying to model earthquakes. Some of them will be described below.

III. DATASETS

When a specific field is researched in terms of machine learning, the first question is where to find data. As for earthquake datasets, various organizations and research institutions are constantly monitoring seismic activity of all over the world. There are some open-source national databases and earthquake catalogs, such as seismicity catalogs of Seismological Institute, National Observatory of Athens (http://www.gein.noa.gr/en/seismicity/earthquake-catalogs, Greece), "Earthquakes of Russia" database of Geological Survey, Russian Academy of Sciences (http://eqru.gsras.ru/, Russia), earthquake list of National Institute of Geophysics and Volcanology (http://cnt.rm.ingv.it/en, Italy) et al. There are also public earthquake catalogs provided by international organizations, which contain earthquake data from all over the world. Some examples are USGS catalog (https://earthquake.usgs.gov/earthquakes/search/), EMSC earthquake database (https://www.emsc-csem.org/) and ANSS Composite catalog by Northern California Earthquake Data Center (http://www.ncedc.org/anss/).

Speaking about the structure of earthquake data, it is usually presented in the form of a table, each record of which corresponds to a certain seismic event. The sets of attributes are different for data published in different catalogs, but the most common ones are:

- time of an event's occurrence;

- geographical coordinates of an epicenter;

- depth of a hypocenter;

- magnitude value, which characterizes the overall "size" of an event and is obtained from measurements of seismic waves recorded by a seismograph;

- magnitude scale used when computing the magnitude value. Several scales have been defined, some of which are easier to compute but have limited applicability, as they cannot satisfactorily measure the strength of the largest events. However, all commonly used scales yield approximately the same values for any given seismic event.

It should be noted that the number of records in all public databases is also different for different countries. It depends not only of seismic activity, but also of development of earthquake monitoring systems in these regions. For example, Japan is known to be the country with the biggest amount of earthquakes recorded. However, according to USGS, the most seismically active place in the world is Indonesia, and Japan has the densest seismic network, which helps them to record more earthquakes [12].

Different level of completeness of earthquake catalogs leads researchers to the need to assess the quality of data they have. There are many different methods of evaluation, one of which is based on Gutenberg-Richter's law [13] – an empirical law that describes the relationship between earthquake magnitude (M) and frequency of occurrence of events (N) for a given region and a time range. It is expressed as:

$$\log_{10} M = a - bN; \quad (1)$$

i.e. the frequency rises exponentially with decreasing magnitude. This relation is remarkably resistant in space and time, so data from complete catalogs should also correspond
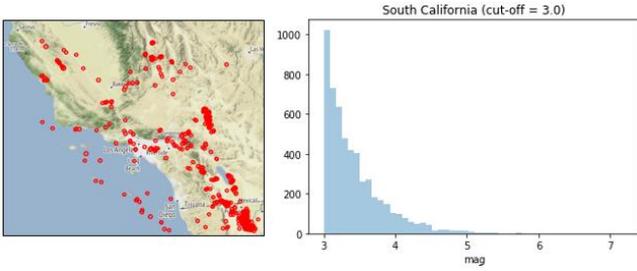
Fig. 1.   The illustration of seismic activity (*left*) and a magnitude distribution plot (*right*) for a region of South California.
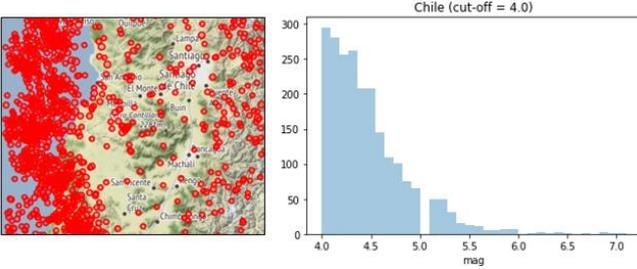

Fig. 2.   The illustration of seismic activity (*left*) and a magnitude distribution plot (*right*) for a region of Chile.
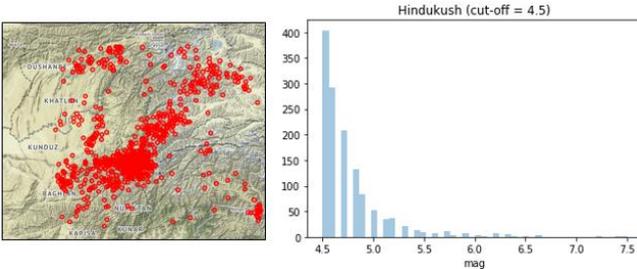

Fig. 3.   The illustration of seismic activity (*left*) and a magnitude distribution plot (*right*) for a region of Hindukush.

to it. The study is performed by plotting magnitude distribution. The point where curve starts to deviate from exponential behavior is chosen as a so-called «cut-off magnitude» (it is stated that some events of a magnitude below this threshold value are missing, so the catalog is incomplete). Therefore, the events of magnitude lower than the cut-off value are removed from the dataset. The illustrations of Gutenberg-Richter law for some frequently studied seismic zones are given in Fig. 1, 2 and 3.

IV. PERFORMANCE MEASURES

In this section, the definition is given for the performance measures that are used in literature to evaluate the prediction models.

The simplest metrics used for quality assessment are:

- *True Positive* (TP): The number of outcomes where the model predicted an earthquake and it actually occurred.

- *False Positive* (FP): The number of outcomes where an earthquake was predicted but did not occur in actual.

- *True Negative* (TN): The number of outcomes where the model predicted no earthquake and there was no earthquake in actual.

- *False Negative* (FN): The number of outcomes where the model predicted no earthquake but it actually occurred.

These measures are summarized in a so-called *confusion matrix* where all possible outcomes are depicted:

TABLE I.   THE CONFUSION MATRIX FOR EARTHQUAKE PREDICTION MODELS

	Occurred	Didn't occur
Was predicted	TP	FP
Wasn't predicted	FN	TN

There are also some other common criteria derived from the above-mentioned ones. Two of the most common statistical measures are *sensitivity* (denoted by S_n, also called POD, that stands for *probability of detection*), or a rate of actual positives predicted correctly, and *specificity* (denoted by S_p), or a rate of actual negatives predicted. They are defined as shown in Eq. 2 and Eq. 3, respectively:

$$S_n = \frac{TP}{TP+FN}; \qquad (2)$$

$$S_P = \frac{TN}{TN+FP} \qquad (3)$$

Two other important evaluation measures are P_1, or a *positive predictive value*, and P_0, or a *negative predictive value*. P_1 is the percentage of true positives out of all positive predictions, as mentioned in Eq. 4, and P_0 is the percentage of true negatives out of all negative predictions of a model, as shown by Eq. 5.:

$$P_1 = \frac{TP}{TP+FP}; \qquad (4)$$

$$P_0 = \frac{TN}{TN+FN} \qquad (5)$$

Some observed papers also define the criteria called *false alarm ratio* (denoted by FAR) and *frequency bias* (FB), as shown by Eq. 6 and Eq. 7, respectively:

$$FAR = \frac{FP}{TP+FP} = 1 - P_1; \qquad (6)$$

$$FB = \frac{TP+FP}{TP+FN} \qquad (7)$$

Accuracy is also computed from four elements of the confusion matrix. It indicates the percentage of number of accurate predictions out of all predictions made by the model. Accuracy is defined as follows:

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \qquad (8)$$

When earthquake prediction problem is formulated as a binary classification task, another performance criteria used are *R score* and *Matthew's correlation score* (denoted by MCC). They are proposed as balanced evaluation measures and are defined as shown in Eq. 9 and Eq. 10, respectively:

$$R = \frac{TP*TN-FP*FN}{(TP+FN)(FP+TN)}; \qquad (9)$$

$$MCC = \frac{TP*TN-FP*FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \qquad (10)$$

Finally, in some papers where regression approach is applied to earthquake prediction, such standard measures as *mean absolute* (MAE) and *relative errors* (RE) are used. They are computed as follows:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|\hat{y}_i - y_i|; \qquad (11)$$

$$RE = \frac{1}{n*\max(y_i)}\sum_{i=1}^{n}|\hat{y}_i - y_i| \qquad (12)$$

V. REVIEW OF EXISTING APPROACHES

This section reviews a number of publications where application of machine learning methods to the task of earthquake prediction on various temporal and spatial intervals have been studied. Due to the fact that, as mentioned above, the processes of earthquake occurrence are considered to be stochastic and non-linear, most recent researches in this area are devoted to the applicability of neural networks to this problem. Another machine learning techniques, specifically, various regression and classification algorithms are also reviewed.

A. E.I. Alves (2006)

Reference [14] was one of the first in proposing artificial neural networks (ANN) for earthquake forecasting. The author, E.I. Alves, was inspired by successful application of similar approaches to the tasks of financial forecasting, which, as he thought, are similar to seismic activity in terms of the chaotic nature of both systems. Financial oscillators such as moving averages (MA), moving averages convergence-divergence (MACD), relative strength index (RSI), etc. were used as input data. The forecast was to indicate time and geographical coordinates of an earthquake within spatial and temporal windows, as well as intensity range on Modified Mercalli Intensity scale (denoted by MMI [15]). The proposed method was tested on data of the region of Azores, Portugal. E. I. Alves stated that it forecasted earthquakes correctly in July 1998 (MMI = 8) and in January 2004 (MMI = 5). However, no statistical measures were computed, so we cannot evaluate the performance of this approach objectively. Though time windows were too wide (the month of the seismic event was forecasted to within ± 5 months), the results were "encouraging" and demonstrated the potential of using neural networks to predict earthquakes.

B. A. Panakkat & H. Adeli (2007), H. Adeli & A. Panakkat (2009)

H. Adeli and A. Panakkat, the authors of [16] and [17], formulate the problem of earthquake prediction as a classification task where the magnitude ranges of the largest seismic event in a pre-defined time window (for instance, 1 month) are the output classes. So, the proposed methods are used to predict the magnitude of the biggest earthquake (within 0.5) in a pre-defined region in the following month.

Reference [16], published in International Neural Systems in 2007, defines eight so-called *seismicity indicators* – mathematically computed features, which can be used to evaluate the seismic potential of a region. These parameters are based of two models of magnitude temporal distribution. The first one is Gutenberg-Richter inverse power law that was described in section "*Datasets*". Another one is characteristic

model, which is based on the fact that some seismic zones exhibit periodic trends in release of seismic energy through large earthquakes. Due to the importance of these indicators for the formation of an approach to the study of the subject of earthquake prediction, their description is provided in Table II.

TABLE II. THE DESCRIPTION OF ADELI & PANAKKAT'S SEISMICITY INDICATORS

Symbol	Description	Mathematical expression
T_θ	Elapsed time	$T = t_n - t_1$
M_{mean}	Mean magnitude	$M_{mean} = \frac{\sum_{i=1}^{n} M_i}{n}$
$dE^{1/2}$	The rate of square root of seismic energy	$dE^{1/2} = \sum E^{1/2}/T$ ($E = 10^{(11.8 + 1.5M)} ergs$)
β	b-value (the slope of Gutenberg-Richter curve)	$\beta = \frac{n\sum(M_i \log N_i) - \sum M_i \sum \log N_i}{(\Sigma M_i)^2 - n\Sigma M_i^2}$
η	Mean square deviation	$\eta = \frac{\sum(\log_{10} N_i - (a - bM_i))^2}{n-1}$ ($b = \beta$; $a = \frac{\Sigma(\log_{10} N_i + bM_i)}{n}$)
ΔM	Magnitude deficit	$\Delta M = M_{max,observed} - M_{max,expected}$ ($M_{max,expected} = \frac{a}{b}$)
μ	Mean time between characteristic events	$\mu = \sum(t_{i\ characterisic})/n_{characteristic}$
c	Coefficient of variation	$c = \frac{STD\ of\ the\ observed\ times}{\mu}$

The logical consequence was paper [17] published in Neural Networks in 2009. In this paper, the authors proposed the architecture of a probabilistic neural network (PNN) as a solution for the same problem that was formulated in [16]. Adeli and Panakkat also used the same set of seismicity indicators as input data for training the network. The model was tested on data for South California seismic zone (33.8-35.4 N° and 114.75-119.25 W°) and yielded good prediction accuracies for events of magnitude 4.5 to 6.0 (R score values between 0.62 and 0.78). However, PNN did not perform satisfactorily for quakes of magnitudes greaten than 6.0, yielding R scores in range from 0.0 to 0.5.

Thus, studies [16] and [17] complement each other: the authors propose using RNN for predicting earthquakes of large magnitude, while PNN may be used for small and moderate earthquakes. The researches of Adeli and Panakkat have laid the foundation for a scientific approach to assessing the potential seismic hazard for different regions: the set of eight seismicity indicators proposed by them was used in various studies by researchers from all over the world.

C. J. Reyes et al.(2013)

In paper [18], published in Applied Soft Computing in 2013, another method for earthquake prediction using ANN is proposed. The system is designed to provide two kinds of predictions: a) the probability that an earthquake larger that a threshold magnitude happens in five days and b) the probability that a seismic event within a pre-defined magnitude range might occur. The input for the proposed predictor was based on b-value from Gutenberg-Richter's law (defined in Table 2); moreover, new seismic parameters were

firstly defined. These parameters are based on Bath's law [19] and Omori-Utsu's law [20], which describe the relations between main shock and aftershocks, based on their magnitude and frequency of occurrence, respectively. Four seismic regions of Chile were analyzed: Talca (35-36 S° and 71-72 W°), Santiago (33-34 S° and 71-72 W°), Pichilemu (34-34.5 S° and 72-72.5 W°) and Valparaíso (32.5-33.5 S° and 71-72 W°). A different feed-forward backpropagation ANN was applied to each area, though they all shared the same architecture. The prototype predicted an earthquake each time when predicted probability was higher than a pre-defined threshold value (the thresholds were adjusted to reduce the number of false alarms). Evaluation of proposed methods was conducted using performance measures computed from TP, TN, FP and FN. Comparative analysis was performed using standard methods of classification such as K nearest neighbors (KNN), support vector machines (SVM) and classification via K-means clustering. Despite the individual setting of parameters, the performance of proposed ANN varied greatly depending on the region: the P_0 values were 17.4% for Talca, 41.7% for Santiago, 86.7% for Pichilemu and 87% for Valparaíso.

D. G. Cortés et al. (2018)

In study [21], which was published in Computers & Geosciences in 2018, an attempt to predict magnitude of the largest seismic event within the next seven days was made.

The problem of earthquake prediction was treated as a regression task: four regressors (generalized linear models, gradient boosting machines, deep learning and random forest) and ensembles for them were applied. Seismicity indicators proposed by Panakkat & Adeli [16] and Reyes et al. [18] were used as input data. The main feature of the study is that the problem was observed in context of big data analytics: a total 1 GB of data processed by means of a cloud-based information were used for training and testing regression models. In order to evaluate the effectiveness of proposed approaches, mean absolute (MAE) and relative (RE) errors were used as performance measures. Besides, due to the specifics of the task, the time spent on training models was also taken into account. The most effective regressor was random forest (RF), yielding a mean absolute error of 0.74 on average. RF was also one of the fastest, taking only 18 minutes to train the regression models on all data. Particularly, the most accurate predictions of RF were made for moderate earthquakes (magnitudes within a range on [4, 7); MAE<=0.26), while regression ensembles performed better on extreme magnitude ranges ([0, 3) and [7, 8)). Based on these results, the authors concluded that using more complex regressor ensembles would improve the accuracy of predictions for quakes of large magnitude.

E. M. Moustra et al. (2011)

The main purpose of study [22] published in Expert Systems and Applications in 2011 was to evaluate the accuracy of ANN for earthquake prediction using different inputs. More specifically, the paper highlights two main areas of research. The first case study concerned prediction of the largest seismic event of the following day using only time series earthquake magnitude data, and the second one concerned the use of so-called Seismic Electric Signals (SES) to predict the magnitude of the next seismic event as well as time lag. For the first case, a feed-forward backpropagation neural network was used. An input file contained maximum magnitude value for each day. The model was trained using

an earthquake catalog for Greece, and performance was evaluated with accuracy rate, which was calculated based on MAE. The average accuracy rate was 80.55% for all events, but only 52.81% for what Moustra et al. considered "outliers" (earthquakes of magnitude greater than 5.2). In order to improve the performance on major quakes, the authors trained the ANN it two phases (at first on outliers, then on all training dataset), and the resulting accuracy rate was 58.02%.

The case study that concerned earthquake prediction using SES consisted of two major parts. It is noteworthy that at the time of the study only 29 samples of SES were recorded and published by VAN team in Greece. Despite this, the authors of [22] tried using an ANN to study the connection between SES and the occurrence of earthquakes. Due to the fact that 29 samples were clearly not enough to train neural networks, Moustra et al. had decided to construct the missing data for the rest of seismic events from the catalog. In first case, SES were generated randomly for all events; in second one the ANN was used to construct missing data using magnitude time series. The accuracy rate of magnitude prediction was slightly more than 60% on the first dataset, and the ANN found no correlation between SES and the time lag. Using data constructed by the ANN improved the performance significantly: the accuracy rates that resulted from the prediction of both magnitude and time lag were 83.56% for magnitude and 92.96% for time lag. The results have led the authors to conclusion that training models on the appropriate data is a key factor that may influence the resulting performance greatly.

F. K. Asim et al. (2017)

In paper [23], which was published in Natural Hazards in 2017, the problem of earthquake prediction is studied as a binary classification task. Predictions were made for events of magnitude greater than or equal to 5.5 on monthly basis. Eight seismicity indicators proposed by Adeli & Panakkat [16] were used as input to different machine learning classifiers. These included recurrent neural network (RNN), pattern recognition neural network (PRNN), random forest (RF) and LPBoost ensemble of decision trees. In addition to the accuracy of predictions, Asim et al. identified such performance measures as sensitivity and specificity, true and false predictive values as the main criteria for comparison of the above-mentioned approaches. The classifiers were used to predict earthquakes in the Hindukush region. LPBoost ensemble tended to take the lead in accuracy with the value of 65%. This classifier also performed better in terms of sensitivity towards earthquake occurrence, yielding 91% of S_n value. The authors also highlighted the result of PRNN, which produced the least false alarms as evidenced by a high level of positive predictive value equal to 71%. Having analyzed the results, the authors stated that every observed system had shown satisfactory results somehow or other.

G. K. Asim et al. (2018)

An earthquake prediction system (EPS) named EP-GPBoost was described in paper [24], which was published in Soil Dynamics and Earthquake Engineering in 2018. This system is a classifier based on a combination of genetic programming (GP) and a boosting algorithm named AdaBoost. An application of these instruments to the problem of earthquake prediction had never studied before this paper. Another novelty of the approach is a methodology of computation and simultaneous usage of seismicity indicators, which is based on idea of obtaining maximum information

about geological properties of observed regions (instead of choosing appropriate parameters for each zone individually). A total of 50 features was calculated, based on such geological concepts as Gutenberg-Richter's law, release of seismic energy, foreshock frequency, etc. Some of these parameters were computed via different approaches (for example, the above-mentioned b-value, which is a slope of a Gutenberg-Richter curve, was computed using two methods, namely, least square regression analysis (as shown in Table 2) and maximum likelihood method). As a result, a system for predicting seismic events of magnitude equal or greater than 5.0 for the next 15 days was proposed. The study of the applicability of EP-GPBoost was performed using data from previously used seismic zones, namely, Chile (32.5–36 S°, 70 –72.5 W°), Hindukush (35-39 N°, 69 –74.6 E°) and Southern California (32 –36.5 N°, 114.75 –121 W°). The experiments have shown outstanding performance in all three observed regions both in terms of low false alarm ratio (the precision values were 74.3%, 80.2% and 84.2% for Hindukush, Chile и Southern California, respectively) and in terms of other metrics considered for evaluation, such as MCC and R score. The best results were obtained for the region of South California (the authors stated that the reason was the quality and completeness of the corresponding earthquake catalog). However, the results of all the regions exhibit improvement when compared to the previous studies [16][18][23].

H. K. Asim et al. (2018)

Reference [25], published in PLOS ONE in 2018, was written by the authors of the previous research. In this paper Asim et al. also used the approach to usage of seismicity indicators proposed in [24]. This time, 60 seismic parameters was computed using various concepts of seismology. Again, some specific features were calculated via different approaches to retain the most complete information about the observed seismic zones. As in their previous research, the authors aim to predict the earthquakes of magnitude equal to or greater than 5.0 for the next 15 days. The proposed system is multistep, unlike previous other predictors proposed in literature which are mainly simple. The system is a combination of different machine learning algorithms, and on each step, one algorithm uses the knowledge obtained through learning of a previous one. Firstly, two-step feature selection is used to choose the most relevant parameters for training a model. Specifically, relevance and redundancy checks are performed (Minimum Redundancy Maximum Relevance criteria, denoted as mRMR, is applied). The resulting set of parameters is passed to a support vector regressor (SVR), and the trend predicted by SVR is then used as a part of input data for a hybrid neural network (HNN). A HNN proposed in [25] is a combination of three different ANNs and EPSO algorithm for weight optimization. The resulting system called SVR-HNN was applied to previously studied regions of Hindikush, Chile and Southern California. The performance was evaluated with such measures as P_0, P_1, S_n, S_p, accuracy, MCC and R score. The results were also compared with ones described in previous researches on these seismic zones. The resulting values of performance measures (for instance, R score increased from 0.27 to 0.58 for Hindukush, from 0.344 to 0.603 for Chile, 0.623 from 0.5107 to 0.623 for Southern California) showed that proposed multistep methodology improved prediction performance in comparison with individual machine learning techniques.

All reviewed papers are summarized in Table III. An analysis of all the above-mentioned works revealed a number of trends in studying the problem of earthquake prediction. Some of these trends and common approaches are described below.

VI. DISCUSSION

This section identifies the main tendencies in earthquake prediction using machine learning techniques and highlights the areas that should be the subjects of further research.

TABLE III.   SUMMARY OF ALL STUDIES REVIEWED

Ref.	Seismic zone	Data sources	ML methods observed	Features	Evaluation techniques
[14]	Azores (Portugal)	-	ANN	Financial oscillators	-
[16]	South California, San-Francisco bay (USA)	SCEC catalog	LMBP; RNN; RBFN	Seismic indicators based on GR law and characteristic distribution	S_n, FAR, FB, R score
[17]	South California (USA)	SCEC catalog	PNN	[16]	S_n, FAR, R score
[18]	Chile	University of Chile's National Service of Seismology	feed-forward ANN	Seismic indicators based on GR law, Bath's law and Omori-Utsu's law	P_0, P_1, S_n, S_p
[21]	California (USA)	ANSS Composite Earthquake Catalog, NCEDS	GLM; GBM; DL	[16], [18]	MAE, RE
[22]	Greece	Seismological Institute, National Observatory of Athens (SINOA), VAN team	ANN	Two approaches to modeling SES for all seismic data	Accuracy rate based on MAE
[23]	Hindukush (Pakistan)	Center for Earthquake Studies (Pakistan), USGS	LMBP-RNN; PRNN; RF; LPBoost	[16]	P_0, P_1, S_n, S_p, accuracy
[24]	Hindikush (Pakistan), Chile, South California (USA)	USGS catalog	GP, AdaBoost	50 seismic indicators used simultaneously	P_0, P_1, S_n, S_p, accuracy, MCC, R score
[25]	Hindikush (Pakistan), Chile, South California (USA)	USGS catalog	mRMR criteria, SVR, HNN, EPSO	60 seismic indicators used simultaneously	P_0, P_1, S_n, S_p, accuracy, MCC, R score

First of all, the definition of an earthquake prediction given by seismologists implies giving the exact definition of time and place of earthquake occurrence as well as its magnitude (as defined in the section "Description of the task"). However, most of the studies observed are focused on wider aim of predicting magnitude for a limited area and temporal range. (The summary of temporal, spatial and magnitude limits used in reviewed papers when formulating the problem are given in Table IV.) That is explained by extreme complexity of the process of earthquake occurrence. urther research in this area should be directed towards attempts to simultaneously predict magnitude, time and place of seismic events' occurrence.

TABLE IV.   THE SUMMARY OF ALL THE APPROACHES TO LIMITING THE PREDICTIONS' ACCURACY USED IN REVIEWED PAPERS

Ref.	Time ranges	Geographical boundaries	Magnitude ranges
[14]	± 5 months	The latitude of an epicenter is predicted to within 1°	-
[16]	1 month	Two pre-defined regions: South California (32-36 N°, 114-120 W°); San-Francisco bay (37.5-40 N°, 116-123.5 W°)	The magnitude is predicted to the nearest 0.5
[17]	15 days	One pre-defined region: South California (33.8-35.4 N°, 114.75-119.25 W°)	The magnitude is predicted to the nearest 0.5
[18]	5 days	The coordinate ranges varied from 0.5° × 0.5° to 1° × 1°	The probability that the magnitude will exceed the threshold value or will be in a pre-defined range is predicted
[21]	7 days	The coordinates of an epicenter are predicted to within a cell of 0.5° × 0.5°	-
[22]	1 days	The pre-defined region of Greece	-
[23]	1 month	The pre-defined region of Hindukush	The occurrence of events of magnitude ≥ 5.5 is predicted
[24]	15 days	Southern California (32-36.5 N°, 114.75-121 W°), Chile (32.5-36 S°, 70-72.5 W°), Hindukush (35-39 N°, 69–74.6 E°)	The occurrence of events of magnitude ≥ 5.5 is predicted
[25]	15 days	Southern California (32-36.5 N°, 114.75-121 W°), Chile (32.5-36 S°, 70-72.5 W°), Hindukush (35-39 N°, 69–74.6 E°)	The occurrence of events of magnitude ≥ 5.5 is predicted

As for data processing, most of the papers reviewed use the approach of feature extraction based on seismic characteristics of a region. As every seismic zone has its unique parameters, it is obvious that these parameters need to be considered for building an exact model. This "personalized" approach is especially noticeable in some of the studies where various zones were observed: the results show that some approaches performed better on one region and worse on the other. There were also researches where different architectures or even methods were applied for modeling different seismic zones because of their differences. In addition, the principles of feature selection and usage are changing over time: in papers published in 2018 a new approach is proposed, which is based on simultaneous use of a large number of seismic indicators for building and training the predicting models.

It is also noteworthy that a number of researches outlines low false alarm generation as an important criterion of performance evaluation. Many authors indicate that earthquake prediction is a delicate issue where false alarms lead to particularly negative consequences, such as economical losses and panic among the civilians, which can be critical because it may cause distrust of the system. Therefore, in some cases we can even sacrifice the sensitivity of a model in favor of reducing a number of false alarms.

Speaking about the performance of proposed models, it is worth noting that it is hard to compare approaches proposed in different papers, because the researchers use different performance measures for assessing the quality of predictors. That is why one cannot objectively state that one model is better than the other is. However, some conclusions can still be made. First of all, the accuracy of predictions as well as other performance measures increase with the research on the field of earthquake prediction (it is noticeable based on the repeatedly studied regions of Southern California, Chile and Hindukush, where similar performance measures have been used). It is also worth noting that in some papers a tendency is observed concerning the decrease of accuracy with increasing magnitude threshold. That is, the larger the earthquake, the harder it is to predict. Given the fact that large earthquakes represent the greatest threat to society, it is necessary to make bigger efforts in the task of predicting earthquakes of high magnitude (equal to or greater than 5.5).

The models proposed in most of the papers reviewed were tested on data for different regions obtained from different earthquake catalogs. We think that this is a major issue. As shown in a number of papers, an approach may perform differently on zones with different seismic properties, and that is another reason why it is near to impossible to compare the methods proposed in different studies. As a solution, we propose to create a «benchmark» dataset, which researchers can use in comparative purposes for different algorithms. The dataset may contain open-source data on seismic zones used in previous studies, such as Chile, Hindukush and Southern California. Besides, we think that it is necessary to complement the dataset with records from other seismic zones from different parts of the world, for instance, Europe and East Asia. We believe that testing the approaches on unified data from regions with different magnitude distributions and other seismological properties will help to carry out a more detailed study of their applicability. The exact geographical boundaries of regions from the proposed «benchmark» dataset and cut-off magnitudes chosen for these regions based of the study of Gutenberg-Richter curves (as described in section "*Datasets*") are listed in Table V. The visualization of seismic activity and magnitude distribution of these regions is shown in Fig 1-5.

TABLE V.   RANGES OF COORDINATE BOUNDARIES OF REGIONS PROPOSED AS PARTS OF THE DATASET

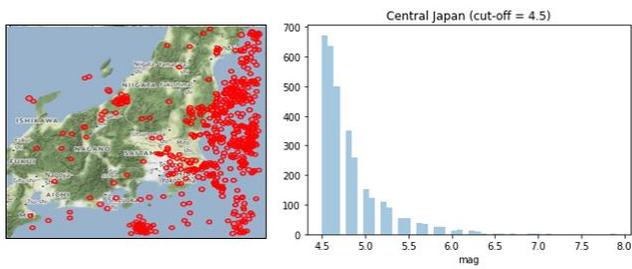| Seismic zones	Latitude range	Longitude range	Cut-off magnitude
Central Japan	34-39 N°	136.5-142 E°	4.5
Chile	32.5-36 S°	70-72.5 W°	4.0
Hindukush	35-39 N°	69-74.6 E°	4.5
Sicily, Italy	36-39 N°	12-16 E°	2.5
South California	32-36.5 N°	114.75-121 W°	3.0

Fig. 4. The illustration of seismic activity (*left*) and a magnitude distribution plot (*right*) for a region of Central Japan.
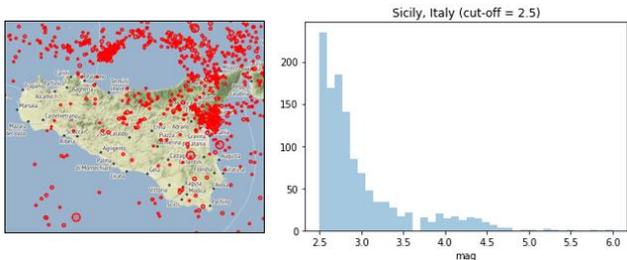

Fig. 5. The illustration of seismic activity (*left*) and a magnitude distribution plot (*right*) for a region of Sicily Island.

VII. CONCLUSION

In this research, the main approaches in application of machine learning methods to a problem of earthquake prediction are observed. The main open-source earthquake catalogs and databases are described. The definition of main metrics used for performance evaluation is given. A detailed review of published works is presented, which highlights the way of development of scientific methods in this area of research. Finally, during the discussion of the results achieved, further directions of research in the field of earthquake prediction are proposed. These are:

- Creating a "benchmark" earthquake dataset, which can be used to assess the quality of various predictor systems. The dataset includes frequently observed seismic zones and seismically active areas of East Asia and Europe, such as Central Japan and Sicily Island. The performance of previously proposed methods can also be evaluated using the «benchmark» dataset.

- Focusing on the most complex and important task of predicting earthquakes of high and extreme magnitudes (equal to or greater than 5.5).

- Making attempts to solve the problem of earthquake prediction in its original form, as determined by earthquake scientists; namely, the simultaneous specification of time, place and magnitude of seismic events with a certain probability.

REFERENCES

[1]   K. Yasuhara, S. Kawagoe, H. Yokoki, and S. Kazama, "Damage from the Great East Japan Earthquake and Tsunami - A quick report," Mitigation and Adaptation Strategies for Global Change, vol. 16(7), pp. 803–818, 2011.

[2]   K. Wang, Q. Chen, S. Sun, and A. Wang, "Predicting the 1975 Haicheng earthquake," Bulletin of the Seismological Society of America, vol. 96, pp. 757–795, 2006.

[3]   R. Geller, "Earthquake prediction: a critical review," Geophysical Journal International, vol. 131, pp. 425–450, 2007.

[4]   R. A. Kerr, "Seismology: Parkfield keeps secrets after a long-awaited quake," Science, vol. 306(5694), pp. 206–207, 2004.

[5]   C. Lomnitz, Fundamentals of Earthquake Prediction. Wiley, New York, NY, 1994.

[6]   B. Rouet-Leduc, C. L. Hulbert, N. Lubbers, K. M. Barros, C. Humphreys, and P. A. Johnson, "Machine learning predicts laboratory earthquakes," Geophysical Research Letters, vol. 44, pp. 9276–9282, 2017.

[7]   P. M. R. DeVries, F. Viégas, M. Wattenberg, and B. Meade, "Deep learning of aftershock patterns following large earthquakes," Nature, vol. 560, pp. 632–634, 2018.

[8]   E. Florido, J. L. Aznarte, A. Morales-Esteban, and F. Martínez Álvarez, "Earthquake magnitude prediction based on artificial neural networks: A survey," Croatian Operational Research Review, vol. 7(2), pp. 159-169, 2016.

[9]   W. Marzocchi, and J. D. Zechar, "Earthquake forecasting and earthquake prediction: different approaches for obtaining the best model," Seismological Research Letters, vol. 82(3), pp. 442–448, 2011.

[10]   D. D. Jackson, "Hypothesis testing and earthquake prediction," Proc. Natl. Aca. Sci. USA, vol. 93, pp. 3772–3775, 1996.

[11]   C. R. Allen, "Responsibilities in earthquake prediction: To the Seismological Society of America," Bulletin of the Seismological Society of America, vol. 66 (6), pp. 2069-2074, 1976.

[12]   "Which country has the most earthquakes?" (Last accessed 24 January 2019) [Online] Available: https://www.usgs.gov/faqs/which-country-has-most-earthquakes

[13]   F. C. Richter, and B. Gutenberg, "Magnitude and energy of earthquakes," Annals of Geophysics, vol. 9, 1956.

[14]   E. I. Alves, "Earthquake forecasting using neural networks: results and future work," Nonlinear Dynamics, vol. 44(1-4), pp. 341-349, 2006.

[15]   C. F. Richter, Elementary Seismology. W. H. Freeman, 1958.

[16]   A. Panakkat, and H. Adeli, "Neural network models for earthquake magnitude prediction using multiple seismicity indicators," International Journal of Neural Systems, vol. 17(1), pp. 13–33, 2007.

[17]   H. Adeli, and A. Panakkat, "A probabilistic neural network for earthquake magnitude prediction," Neural Networks, vol. 22(7), pp. 1018–1024, 2009.

[18]   F. Martínez-Álvarez, A. Morales-Esteban, and J. Reyes, "Neural networks to predict earthquakes in Chile," Applied Soft Computing, vol. 13, pp. 1314-1328, 2013.

[19]   M. Bath, "Lateral inhomogeneities in the upper mantle," Tectonophysics, vol. 2, pp. 483–514, 1965.

[20]   T. Utsu, "A statistical study of the occurrence of aftershocks," Geophysical Magazine, vol. 30, pp. 521–605, 1961.

[21]   G. Cortés, A. Morales-Esteban, X. Shang, and F. Martínez-Álvarez, "Earthquake Prediction in California Using Regression Algorithms and Cloud-based Big Data Infrastructure," Computers & Geosciences, vol. 115, pp. 198-210, 2018.

[22]   M. Moustra, M. Avraamides, C. Christodoulou, "Artificial neural networks for earthquake prediction using time series magnitude data or Seismic Electric Signals," Expert Systems & Applications, vol. 38, pp. 15032-15039, 2011.

[23]   K. Asim, F. Martínez-Álvarez, A. Basit, and T. Iqbal, "Earthquake magnitude prediction in Hindukush region using machine learning techniques," Natural Hazards, vol. 85, pp. 471-486, 2017.

[24]   K. Asim, A. Idris, T. Iqbal, and F. Martínez-Álvarez, "Seismic indicators based earthquake predictor system using Genetic Programming and AdaBoost classification," Soil Dynamics and Earthquake Engineering, vol. 111, pp. 1–7, 2018.

[25]   K. Asim, A. Idris, T. Iqbal, and F. Martínez-Álvarez, "Earthquake prediction model using support vector regressor and hybrid neural networks," PLOS ONE, vol. 13, 2018.