IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received April 10, 2021, accepted April 24, 2021, date of publication April 27, 2021, date of current version May 6, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3076119

Comparative Performance Assessments of Machine-Learning Methods for Artificial Seismic Sources Discrimination

MOHAMED S. ABDALZAHER[1], (Member, IEEE), SAYED S. R. MOUSTAFA[1], MOHAMMED ABD-ELNABY[2], AND MOHAMED ELWEKEIL[3]

[1]Seismology Department, National Research Institute of Astronomy and Geophysics (NRIAG), Cairo 11421, Egypt
[2]Department of Computer Engineering, College of Computers and Information Technology, Taif University, Taif 21944, Saudi Arabia
[3]Department of Electronics and Electrical Communications Engineering, Faculty of Electronic Engineering, Menoufia University, Menouf 32952, Egypt

Corresponding author: Mohamed S. Abdalzaher (msabdalzaher@nriag.sci.eg)

This work was funded and supported by the Taif University Researchers Supporting Project Number (TURSP-2020/147), Taif University, Taif, Saudi Arabia. This work was also supported in part by the Egyptian National Seismic Network (ENSN) Lab, National Research Institute of Astronomy and Geophysics (NRIAG).

ABSTRACT Mankind is vulnerable to artificial seismic sources and accompanying explosions' consequences. Recently, seismicity catalog contamination is among the main problems faced by seismologists. Since identifying artificial seismic sources is the first and always challenging stage, it is imperative to develop an automated control system that will discriminate tectonic from non-tectonic events. Detection and removal of the artificial seismic sources have become urgent. Early treatments and cleaning of contaminated seismicity catalogs are crucial to assist in accurate seismic hazard identification and enhance the planning of future urban developments. With the advancement of machine learning (ML) techniques, artificial seismic source detection accuracy has been improved. Today, there are different kinds of methods, ML techniques, and diverse processes like knowledge discovery are developed for discriminating artificial seismic sources and earthquakes. ML techniques offer various probabilistic and statistical methods that allow intelligent systems to learn from reoccurring experiences to detect and identify patterns from a dataset. This study aims to build an automated system that is able to detect the existence of artificial seismic sources in seismicity catalogs. More concretely, we classify seismic activity reports into two classes using classical and ensemble ML algorithms. Classical seismicity parameters or features are supplied to linear and nonlinear ML classifiers. The proposed scheme based on the four features (Latitude, Longitude, depth, and Magnitude) can enhance the performance. To assure the enhanced performance, we have examined the proposed scheme by both the accuracy of each model, ROC curves, Precision-Recall, and Calibration. The obtained results prove that the ensemble learning algorithms exhibit better results compared to other classical ML algorithms by having 98.14% testing accuracy.

INDEX TERMS Machine learning, artificial seismic sources, seismicity contamination, seismic hazard.

I. INTRODUCTION

Northern California contains many artificial seismic sources, where explosives such as dynamite are predominantly used in order to obtain raw materials that are used in the cement industry. These explosives generate artificial vibrations that are obviously recorded by the Northern California Earthquake Data Center (NCEDC), which has established to record earthquakes even at a very small scale. These blasts can contaminate the earthquake catalog and lead to an incorrect assessment of seismic risk and hazard analysis. It is necessary to use a clean earthquake catalog in vital socioeconomic studies. Thus, it is very important to exclude the blasts and decontaminate the local earthquake catalog.

Existing seismological networks record heterogeneous signals generated not only by natural earthquakes but also by artificial man-made sources. Specifically, seismometers at seismic stations record all types of earth vibrations in the

The associate editor coordinating the review of this manuscript and approving it for publication was Mostafa M. Fouda.

region without the ability to clarify their origin. Industrial artificial seismic sources are commonly recorded due to their low energy content in areas where the detection threshold is low. The inclusion of these events in earthquake catalogs can affect seismic hazard studies for the mapped areas in terms of both space and time distribution. These contaminations represent a source of error and falsify seismicity rates, frequency-magnitude distributions, and microseismicity analysis in general and could be misinterpreted as a change in the natural seismic activity [1]. Considering that misidentified artificial seismic events, such as artificial seismic sources and underground nuclear tests, can lead to erroneous analyses, the classification of the signal's source should be performed as a preliminary work prior to seismic signal processing and analysis.

It is a fact that the discrimination between earthquakes and artificial seismicity is a serious issue in seismology as raised by [2], [3]. Currently, seismologists manually classify artificial seismic sources based on visual inspection of the waveforms which is a slow procedure that increases the time to production of scientific results. From this point of view, the detection, mapping, and removal of artificial seismic sources are a preliminary and important step in the analysis based on seismicity catalogs [4]–[6]. Man-made contamination has a low magnitude, so the magnitude cut-off of the seismicity data set [7] could be a solution. However, this implies a lack of insufficient data. Furthermore, this introduces uncertainties in the classification procedure. Both of these problems can be mitigated by using an automated approach. Today, there are different kinds of ML techniques, and knowledge discovery processes are developed for discriminating artificial seismic sources and earthquakes. ML techniques offer various probabilistic and statistical methods that allow intelligent systems to learn from reoccurring experiences to detect and identify patterns from a dataset.

The discrimination between natural earthquakes and man-made explosions has become a predominant problem in seismic activity observation that consumes many efforts. Moreover, the discrimination of small magnitude events is more complex. In the literature context, several waveform-based techniques have been proposed such as ratios between amplitudes of different seismic phases, analysis of coda waves, and detection of ripple-firing spectral modulation [2], [3], [8]–[10]. Fundamentally, seismic discrimination is manually executed by either visually inspecting the natural and man-made events records or by numerical computations of the records' characteristics. However, this requires a great amount of work and time from earthquake analysts. Therefore, several earthquake discrimination errors occur.

Interestingly, the discrimination between the earthquakes and artificial seismic sources using simple methods, i.e., satellite images, occurrence times, is probably not efficient if the artificial seismic sources are located along active fault zones [1], [11]. Consequently, the discrimination can be done utilizing variant techniques for both the time, and frequency domains, e.g., the peak amplitude ratio of the seismic phases [12], [13], the spectral ratio of seismic phases, or average amplitude in low- and high-frequency bands for a specific phase [9], [14]–[16].

Apart from seismological procedures, ML can provide a lot of support in discriminating tectonic seismological events from non-tectonic. Supervised ML algorithms need annotated data for classifying the type of seismic events into tectonic or non-tectonic categories [17]. In the last decade, massive progress has been done in resolving crucial issues in several projects. Recent explosions have attracted many researchers around the globe to solve this problem. Data provided by the NCEDC in the form of seismic activity catalogs [18]. Thus, ML models can be built to classify a seismic activity into an explosion or not. On the other hand, in this paper, we use supervised ML techniques for classifying the seismic activity into two different categories-tectonic and non-tectonic events. The seismic discrimination method using the ML classifier can shorten the workload and increase the reliability of the classification results [9]. A considerable number of seismic signal discrimination methods based on statistical ML have been proposed [19]–[24]. Although, several efforts have been exerted for classifying the seismic activity into two different categories-tectonic and non-tectonic events, a more adaptive and intelligent solution is still desired. This paper essentially presents a ML scheme that relies only on minimal number of features (Latitude, Longitude, depth, and Magnitude) for discriminating the earthquakes and the artificial seismic sources reflecting the simplicity and resilience of the model. The motivation of using these four features is that they can be supported by any seismic network. To the best of our knowledge, no similar approach has been proposed in the literature.

The contributions of this research work are five folds:

- The proposed scheme employs several ML algorithms to discriminate artificial seismic sources from earthquakes with high accuracy. Thus, it can alleviate the impact of improper estimation of seismic hazards. This can assist in early earthquakes' detection and alleviating their consequences leading to a dramatic reduction in the earthquake risk.
- We have proposed a novel classification scheme that relies on only four features (Latitude, Longitude, depth, and magnitude) collected from the utilized catalogs. To the best of our knowledge, these specific four features have not been employed before to solve the considered problem.
- The research work presents a comprehensive overview of the applicability of ML techniques in seismicity analysis by applying learning algorithms on the seismicity catalog dataset.
- We present a taxonomy of the candidate linear and nonlinear ML models that can be utilized in solving the considered problem.
- The obtained results show the effectiveness of utilizing ML for discriminating the earthquake catalogs contaminated by artificial seismic sources. Moreover, this

discrimination is examined by comparing the obtained results of the common linear and nonlinear ML models showing the outperforming technique. Besides, our scheme is compared to deep learning (DL) showing outperforming results using ML, which ensure the non-suitability of DL for such designated problem that relies on only four features. More concretely, the accuracy of the proposed scheme is examined by ROC curves, Precision-Recall, and Calibration.

The rest of this paper is organized as follows. The employed ML classifiers are discussed in Section II. The adopted evaluation methodology is presented in Section III. The framework of the experimental setup is illustrated in Section IV. The obtained results and discussions are portrayed in Section V. Finally, Section VI concludes the paper.

II. EMPLOYED MACHINE LEARNING CLASSIFIERS

Normally, a human is needed to go through data and decide what the classification of an object is. With ML the aim is to make this operation as automated as possible using suitable ML algorithms. We develop several methods for classification in this research. In what follows, we present a brief theory on how each of the utilized classification methods works. In other words, we here present a taxonomy of the common linear and nonlinear ML models that have been employed in our scheme.

A. LINEAR CLASSIFICATION MODELS

1) LOGISTIC REGRESSION

Logistic regression (LR) is utilized in statistical analysis on a dataset with often several independent variables leading to a binary decision. More particularly, LR models the posterior probability of K feature classes to fit data over a logit function yielding a binary output $\in \{0, 1\}$. When there are more than two outcomes, it is classified by multinomial logistic regression. The first is to classify the inputs as class 0 or 1. LR evaluates the observation probability that belongs to class 1. Then, it depends on a logit function for the probabilities to classify the elements into the two target classes (0 and 1). Afterwards, it defines a thresholding level for values between 0 and 1, to ensure that all the values $\in \{0, 1\}$ [25].

2) LINEAR AND QUADRATIC DISCRIMINANT ANALYSIS

Linear Discriminant Analysis (LDA) is developed to resolve several issues confronting the LR. In the case of sufficient splitting of classes, the LR model is not suitable for parameter estimation. On the other hand, if the sample size is quite small and the predictors' distribution is normal in each of the classes, the LDA model is more applicable than the LR model. Before utilizing LDA, an assumption about the data to be processed is desired, in which each of the predictors should be normally distributed. Conversely, LDA can be insufficient in other cases, when the data is easily defined as being higher or lower than a linear assumption. In these cases, also the Quadratic Discriminant Analysis (QDA) can be sufficient but relying on a non-linear assumption [26].

3) LINEAR SUPPORT VECTOR MACHINE

The main aim of the support vector machine (SVM) is to find an optimal hyperplane (a line for two feature classes) for K-feature classes (making a K-dimensional space). The K-dimensional space classifies the input data features to the target classes aiming at achieving the optimal hyperplane. This hyperplane is attained when having the maximum distance separating the class data points. SVMs represent the closest data points to the hyperplane, where SVMs also influence the hyperplane behavior. The linear SVM (LSVM) starts from the output of the linear function, such as the LRLR method, then produces values $\in [-1, 1]$. Afterward, the SVM maximizes the boundary between the data points and hyperplane using a cost function with a regularization parameter for balancing the loss and the boundary maximization. Then, the cost function is minimized considering the weights updating for calculating the gradient. If misclassification occurred, the gradient would be updated by adapting the loss of the regularization parameter [27], [28].

4) RIDGE CLASSIFIER

The Ridge Classifier is generally utilized as a regression method, where it maps the label data $\in [-1, 1]$. Accordingly, it solves the problem based on a regression philosophy. Then, the highest value in prediction is mapped to the target class but multi-output regression is employed for the case of multiclass [29], [30].

5) GAUSSIAN NAIVE BAYES

Generally speaking, the Naive Bayes (NB) classifier is nonlinear. However, if the likelihood factors depend on exponential families, the NB classifier is handled as a linear classifier. A Gaussian algorithm is a special form of the NB algorithm that is employed for the case of features of continuous values. More particularly, the features are supposed to pursue a Gaussian distribution [31]. The Gaussian NB (GNB) is considered a probabilistic model, which predicts the classes based on the probability that each class belongs to. The probability can be given as:

$$p(x = v|C_k) = \frac{1}{\sqrt{2\pi\sigma_k^2}} \exp - \frac{(v - \mu_k)^2}{2\sigma_k^2}, \tag{1}$$

where x is a continuous data input attribute, σ is the variance, v is the probability density, C_k is the class, μ is the mean.

B. NON-LINEAR CLASSIFICATION MODELS

1) ADABOOST

The AdaBoost (AB) adopts adaptive boosting to apply successive weak classifiers for recurrently modifying data sets by gathering them to a better classifier based on a weighted majority vote. The AB algorithm involves more weights in entities that are hard to classify and less on the ones that are easy to hand. In the former stage, AB initializes the observation weights [32]. Then, the classifier is fit to the training data using the weights followed by calculating the weighted

error rate, where another weight is provided to the classifier to determine the final decision of the classifier. Then, the new classifier is examined on the previous classifier, such that the new tree hopefully has a better prediction. Therefore, the generated model is the combination of the two trees. This procedure is recursively done many times of a given number of iterations [33].

2) GRADIENT, LIGHT GRADIENT, AND EXTREME GRADIENT BOOSTING CLASSIFIERS

First, the Gradient boosting (GB) is a model that switches the weak classifiers into stronger ones. Unlike the AB, in the GB the gradient determines the shortcomings of the weak classifiers. Moreover, the GB uses high weight data points in the loss function. Accordingly, GB allows generic and specified cost function for optimization that is suitable for the designated classification problem [33]. Second, the Light Gradient boosting (LGB) is an adapted framework of GB, which utilizes the tree-learning concept. In other words, LGB is developed to be a distributed model that is more efficient than the GB [34]. Third, the Extreme Gradient Boosting (XGB) is an optimized version of the GB library with high efficiency, flexibility, and portability. XGB supports a rapid and efficient parallel tree boosting [35]. Furthermore, the XGB considers the potential loss for the possible splits to make a new branch by looking at distributions of the features for all data points in a leaf. This is used to decrease the space of possible feature splits search. This method involves many hyperparameters for optimal tuning.

3) RANDOM FOREST, DECISION TREE, AND EXTRA TREES

First, the random forest (RF) is among the ensemble methods that rely on the tree concept. It consists of parallel learners that are employed to simultaneously alleviate the bias and variance [36]. Generally speaking, the RF is utilized as a probabilistic predictor in classification problems, which is defined as:

$$\hat{\boldsymbol{Y}}(\mathbf{x}) = \frac{1}{NT}\sum_{i=1}^{NT} R_i(\mathbf{x}), \tag{2}$$

where $\mathbf{x}$ is the vectored input, NT is the number of trees, and $R_i(\mathbf{x})$ is the ith regression tree.

Second, the extra-trees (ET) model is a developed version of the random forest, which is less probable to overfit. It can randomly select the best features among the input dataset to assist the estimators to achieve better performance [37]. Third, the decision tree (DT) model suitable for both classification and regression problems, which is able to handle complex problems by dividing them into multiple simpler ones aiming at easier solution fulfillment [38], [39].

4) K-NEAREST NEIGHBORS

The k-nearest neighbors (KNN) algorithm is a decision-boundary-based classification algorithm that classifies an input to the majority class of its k nearest neighbors in space [40]. The tunable hyperparameters for this algorithm include k, the number of neighbors used for classifying any given input, and the distance metric used for determining the closest neighbors. The minimal KNN is expected when k = 1 because the model is prone to becoming heavily reliant on noise or outliers. In addition, variants of KNN may weigh the votes of training-set observations with their cosine similarity relative to their input.

5) CATBOOST CLASSIFIER

CatBoost (CB) relies on the oblivious trees, which is a specific kind of depth-first expansion. More concretely, it utilizes a vectorized exemplification of the tree for which a binary splitting technique is adopted in each level. Accordingly, it leads to fast convergence and rapid evaluation. On the contrary, the CATBoost is not beneficial to be used with low false-positive rates [41].

III. EVALUATION METHODOLOGY

For classification, the model performance relies on the accuracy score, which is represented by the correct predictions divided by the total number of predictions:

$$\text{Accuracy} = \frac{\sum_{i=1}^{n} I(t_i = y_i)}{n} \tag{3}$$

Here n is the total number of predictions, t_i is the predicted target, y_i is the class target and I is an indicator function. The optimal accuracy score is 1, which means that the prediction fits the data perfectly. With the *cross_val_score* function in the Scikit-Learn library, we also calculate the test cross-validation score with 10 folds with the test data for the chosen classifier.

Another score function is the Cohen Kappa score, which is the accuracy that occurs through random predictions. It measures how well the classifier actually performs, and the best score is 1 as for the accuracy score. The Kappa score is calculated as ([42])

$$\kappa = \frac{\text{Observed Accuracy} - \text{Expected Accuracy}}{1 - \text{Expected Accuracy}}. \tag{4}$$

Moreover, both precision-recall and F1-score metrics can also be used for classification evaluation. They can be defined as:

$$precision = \frac{TP}{TP + FP}, \tag{5}$$

$$recall = \frac{TP}{TP + FN}, \tag{6}$$

$$F1 - score = 2 \times \left(\frac{precision \times recall}{precision + recall}\right). \tag{7}$$

where TP is the true positive decision, FP represents the false positive decision, and FN denotes a false negative decision.

Another plot that we can compute with these prediction results is the Receiver Operating Characteristic (ROC) curve. This plots the true positive rate (sensitivity) as a function of the false positive rate (specificity) for the class 0, class 1,

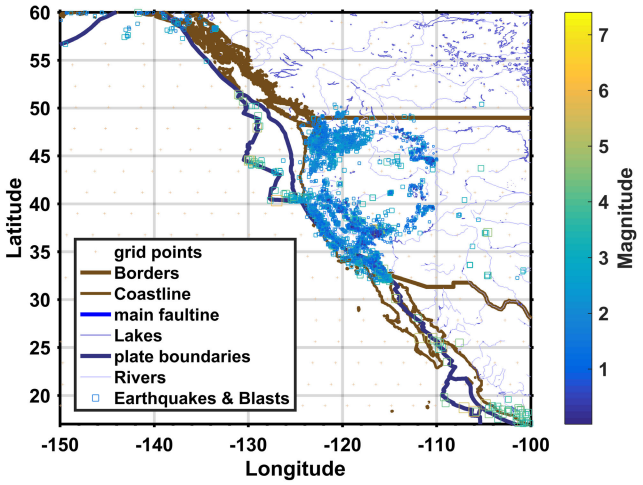


FIGURE 1. Spatial distribution of collected data download from the NCEDC repository.

micro-average, and macro-average for the classifier. The earlier the true positive rate reaches 1.0, the better the accuracy of the classifier. By plotting the cumulative gain of the class targets as functions of the percentage of samples, we can see how effective the classifier is at predicting the class targets as the percentage of the samples vary [43].

IV. EXPERIMENT SETUP

Feature and classifier selection are both crucial for classification using ML; analyzing the key properties of a dataset with the model best suited for that domain tends to lead to better results. Classical ML classifiers require numerical values representing observations of each class.

A. DATA COLLECTION

Based on the Northern California Seismic Network (NCSN), the data is observed by 13 seismic networks (See Figure 1). Since 1984, on average, these networks record ≈20,000 earthquakes per year with ≈1 million seismograms that are archived every year. The main part of the earthquakes that exist at diverse and complex tectonic settings such as the San Andreas Fault system, demonstrating the margin between the Pacific and North American plate, the volcanic region of Long Valley Caldera, and the Mendocino Triple Junction at the intersection of the Gorda, North American, and Pacific plate (See Figure 1). Besides, a wide range of anthropogenic earthquakes are involved in the Geysers Geothermal Field in terms of the geothermal activities [44].

Earthquakes recorded by the NCSN are recorded using by an event-by-event basis. Seismicity data utilized in our research are collected from the NCEDC [18]. This center is the main repository and distribution center for various digital data types of earthquakes in central and northern California. Collected data contain the earthquake parameters catalog of the NCSN from 1966 to the present. The following table (Table 1) gives a list of descriptive statistics of collected earthquakes and artificial seismic sources data.

Here, we benefit from an extensive and regularly archived digital seismogram and parameter data from the NCEDC in most countries over the recent decades, and from the rise in storage and processing power over the past couple of years. We use the densely scattered reported events throughout much of California over recent years. The spatial distribution of collected data download from the NCEDC repository is shown in Figure 1.

B. RELEVANT DATASET PREPROCESSING

In this research, we will study a binary classification problem involving the discrimination of tectonic and non-tectonic seismic events collected and downloaded from the NCEDC repository. The data file contains feature data about possible seismic event type candidates and is classified as tectonic (positive, 1) or non-tectonic (negative, 0). Each event is stored in a separate row and contains four continuous variables and the class label (0 or 1). The first two variables defend the spatial location of the event in terms of latitude and longitude, while the other two variables correspond to the selected event magnitude and depth. Their descriptive statistics can be seen in Table 1. The four continuous feature variables are set as the design matrix $\mathbf{X}$, and the class variable is the target vector $\mathbf{y}$. The data set looks to be complete without any NAN values.

Real-world seismic activity data is commonly chaotic. In fact, refining data by correcting data structure and removing noise is a prerequisite before the data analysis. In this study, we have adapted several preprocessing steps. First, the dataset is normalized to the corresponding dimensions to span the same range. This results in higher accuracy during data processing, as the inter-dimensional variance is accounted for. For our data, each dimension has been divided by its standard deviation simultaneously when it has been zero-centered.

From Table 1, it is clear that the downloaded data is imbalanced as the number of earthquakes exceeds the total number of explosions. Imbalanced classifications pose a challenge for predictive modeling as most of the ML algorithms used for classification were designed around the assumption of an equal number of examples for each class. This results in models that have poor predictive performance, specifically for the minority class. This is a problem because typically, the minority class is more important, and therefore the problem is more sensitive to classification errors for the minority class than the majority class. In fact, the data of the area of interest that we filtered from the whole catalog, is balanced. In other words, the number of earthquakes and artificial seismic sources is relative as shown in Table 2.

To determine which features may be the most effective for classification, the following plots are constructed to visualize the relationship between select features of the earthquakes and artificial seismic sources. Spatial distribution of selected events is given in Figure 2. The lower part of this figure depicts the relationship between depth and magnitude variables along with histograms of each variable in the margins. Not only you can see the relationships between the two

TABLE 1. Descriptive statistics of collected earthquakes and artificial seismic sources data.

Variable	Even Type	Count	Mean	Minimum	Q1	Median	Q3	Maximum
Latitude	Earthquakes	957378	37.985	34.5	37.398	37.655	38.817	41.999
	artificial seismic sources	19898	37.755	34.935	37.314	37.501	38.072	41.917
Longitude	Earthquakes	957378	-121.19	-126	-122.79	-121.53	-118.94	-117.76
	artificial seismic sources	19898	-121.59	-124.39	-122.14	-122.1	-121.02	-117.76
Depth	Earthquakes	957378	4.628	-3.135	1.674	3.541	6.27	120.335
	artificial seismic sources	19898	0.2124	-2.886	-0.436	-0.313	-0.155	73.403
Magnitude	Earthquakes	957378	1.2364	0	0.76	1.12	1.62	7.2
	artificial seismic sources	19898	1.7133	0.07	1.37	1.7	2.01	4.58
No.Stations	Earthquakes	957378	16.704	0	8	12	21	400
	artificial seismic sources	19898	13.215	0	7	11	17	112
Gap	Earthquakes	957378	122.4	0	74	107	155	360
	artificial seismic sources	19898	119.1	25	73	107	151	344
Distance	Earthquakes	957378	6.329	0	1	3	7	303
	artificial seismic sources	19898	6.507	0	2	4	7	175
RMS	Earthquakes	957378	0.0553	0	0.02	0.04	0.07	33.66
	artificial seismic sources	19898	0.1034	0	0.05	0.08	0.12	18.28

TABLE 2. Descriptive statistics of selected data for ML classification.

Variable	Event Type	Count	Mean	Minimum	Q1	Median	Q3	Maximum
Latitude	Earthquake	19006	37.911	35.202	36.64	37.426	38.794	49.96
	Explosion	15153	40.559	35.202	37.325	37.988	46.295	50.382
Longitude	Earthquake	19006	-120.06	-124.57	-122.22	-120.44	-117.95	-115.99
	Explosion	15153	-121.75	-124.57	-122.33	-122.1	-121.57	-115.99
Depth	Earthquake	19006	5.9087	-2.53	2.37	5	8.0125	60.04
	Explosion	15153	0.7055	-2.89	-0.36	-0.22	0.34	78.34
Magnitude	Earthquake	19006	1.3317	0	1	1.26	1.65	5.2
	Explosion	15153	1.7367	0.07	1.35	1.67	2.01	5.7

variables, but also how they are individually distributed. The proposed scheme of utilizing the linear and nonlinear ML models relying on both de-clustered and non-de-clustered catalog is shown in Figure 3.

V. RESULTS AND DISCUSSION

The analytical approach to solving the problem of seismic events classification would primarily contain three steps, feature extraction, building a statistical model of the extracted features, and cross-comparison of the designed model with principal indicators of dominant classes as outlined in the following subsections. However, from collected events magnitudes given in Table 2, it is expected that the utilized catalog contains seismic aftershocks. Aftershocks earthquakes are the successive ones to the largest shock of an earthquake sequence. The magnitudes of these aftershocks are smaller than the accompanying main-shock. Generally, aftershocks could appear within weeks, months, or even years after the main-shock. It is worth mentioning that the aftershocks number is directly proportional to the magnitude of the main-shock.

In our analysis, we have adopted three scenarios. First, we discriminate between three classes (0, 1, and 2) representing the aftershocks, main-shock, and artificial seismic sources, respectively. Second, only two classes (1 and 2) are utilized, one is for only the main-shocks (after de-clustering execution), and the other is for the artificial seismic sources. Similarly, in the third scenario, two classes are also utilized but the first class (1) generally represents the earthquakes (main-shocks and aftershock), which means no de-clustering is done, and the second class (2) denotes the artificial seismic sources. We have implemented classification algorithms using Python codes relying on their respective functions from the Scikit-Learn library. It is worth mentioning that we have used a split size of 80% and 20% to the training and test sets, respectively. Besides, an empirical process where extensive experiments have been done to conclude the set of best configuration parameters leading to the best models

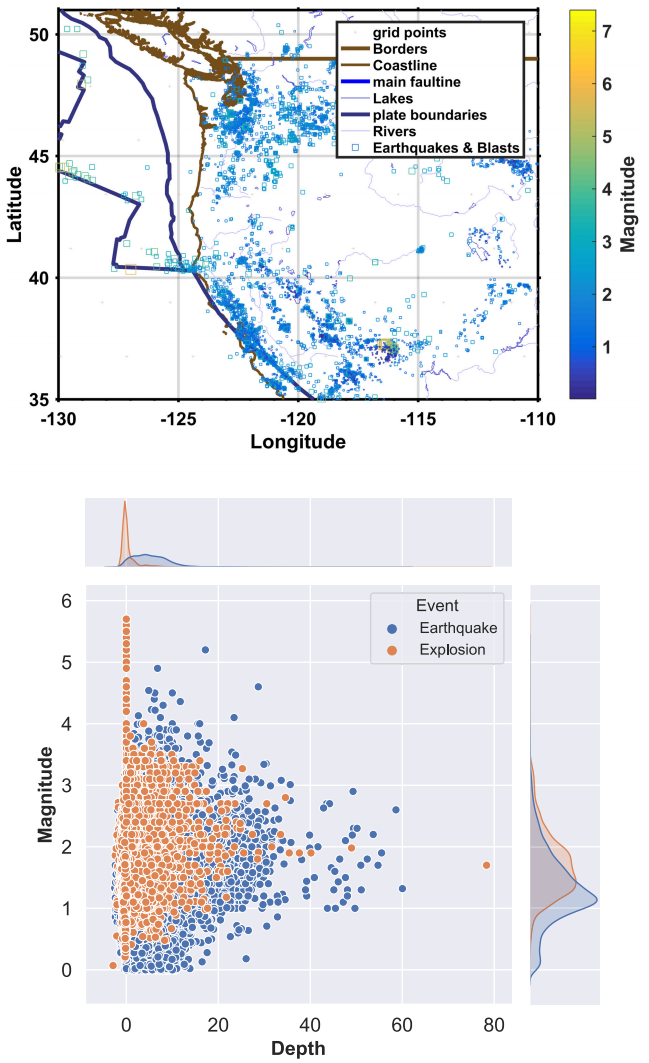


FIGURE 2. Spatial distribution of selected events (upper panel) and marginal plot between depth and magnitude variables (lower panel).

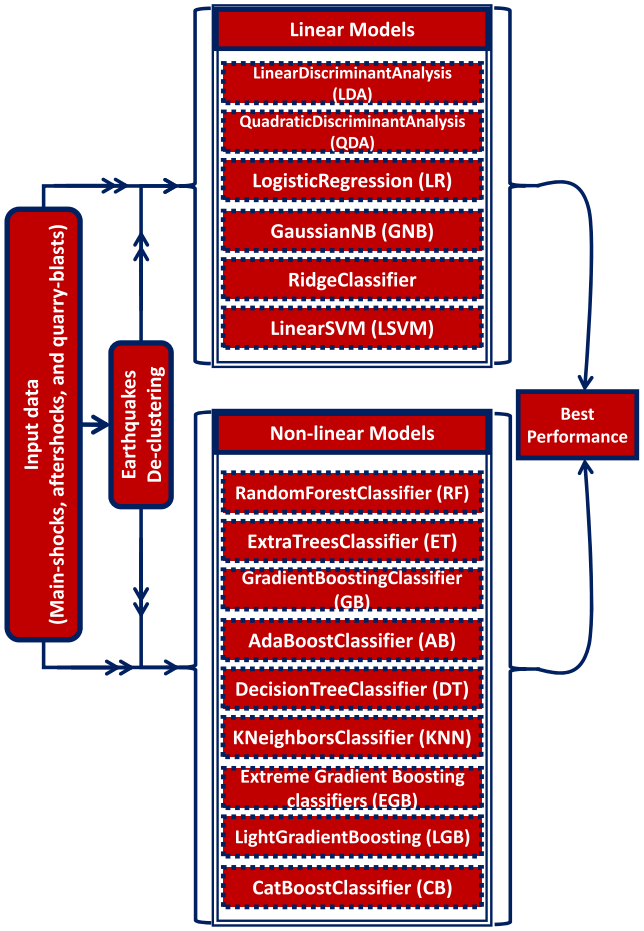


FIGURE 3. Flowchart of the workflow utilized.

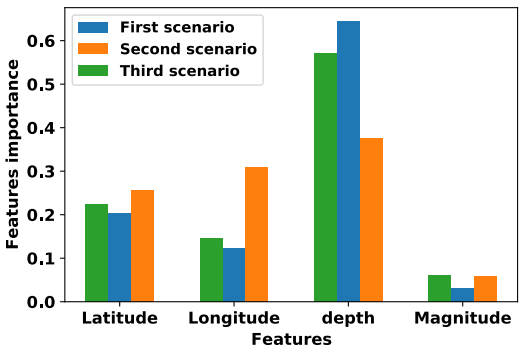


FIGURE 4. Features importance percentage of best models in the three scenarios.

in the three scenarios, which are listed in Table 3. Figure 4 depicts the importance percentage of the utilized features (Latitude, Longitude, depth, and Magnitude) of best models (XGB, ET, and XGB) in the first, second, and third scenarios, respectively.

Figure 5 compares between the obtained accuracy of each of the developed scenarios, i.e., three classes (main-shocks, aftershocks, and artificial seismic sources), two classes with de-clustering (main-shocks and artificial seismic sources), and two classes without de-clustering (earthquakes including both main-shocks/aftershocks and artificial seismic sources). Indeed, we have executed extensive analysis relying on six linear ML models and nine nonlinear ones applied for the above-mentioned three scenarios. It is worth mentioning that almost always the nonlinear models outperform the linear ones. It is clearly shown that the best performance is obtained using the third scenario of earthquakes and artificial seismic sources with 98.14% achieved by one of the nonlinear models, namely, ''ExtremeGradientBoosting''. Meanwhile, the best linear model ''QuadraticDiscriminantAnalysis'' attains accuracy with only 87.17%. The classification accuracy using the second scenario comes in the second rank with a tiny less accuracy by 98.1% relying also on a nonlinear approach ''ExtraTreesClassifier''. Similarly, the linear model performance lags the corresponding nonlinear ones in the second scenario. Finally, the third classification accuracy is attained by the first scenario, in which the best performance is obtained by ''ExtremeGradientBoosting'' model with 95.26%.

TABLE 3. Configuration parameters of the best models in the first, second, and third scenarios.

Parameters description of first and third models (XGB)	First model parameters values	Third model parameters values	Parameters description of second model (ET)	Second model parameters values
Base score	0.5	0.5	Using bootstrap sample	False
Booster	'gbtree'	'gbtree'	Cost-complexity pruning	0
Subsample columns ratio per level	1	1	Class weight	None
Subsample columns ratio per split	1	1	Split quality measuring function	'gini'
Subsample ratio of columns constructing each tree	1	1	Maximum depth	None
Minimum loss reduction	0	0	Maximum features	'auto'
Importance type	'gain'	'gain'	Maximum leaf nodes	None
Learning rate	0.300000012	0.300000012	Maximum samples	None
Maximum delta step	0	0	Minimum impurity decrease	0
Maximum depth	6	6	Minimum impurity split	None
Minimum child weight	1	1	Minimum number of samples per leaf	1
Number of estimators	100	100	Minimum number of split samples per node	2
Number of parallel threads	-1	-1	Minimum weighted fraction per leaf	0
Number of parallel trees	1	1	Number of estimators	100
Objective	'multi:softprob'	'binary:logistic'	Number of parallel threads	-1
Random state	123	123	Using out-of-bag samples for estimation	False
L1 Regularization term on weights	0	0	Random state	123
L2 Regularization term on weights	1	1	Verbosity degree	0
Balancing of positive and negative weights	None	1	Reusing the solution of the previous call	False
Subsample ratio of the training instance	1	1		
Tree method	'auto'	'auto'		
Validate parameters	1	1		
Verbosity degree	0	0		

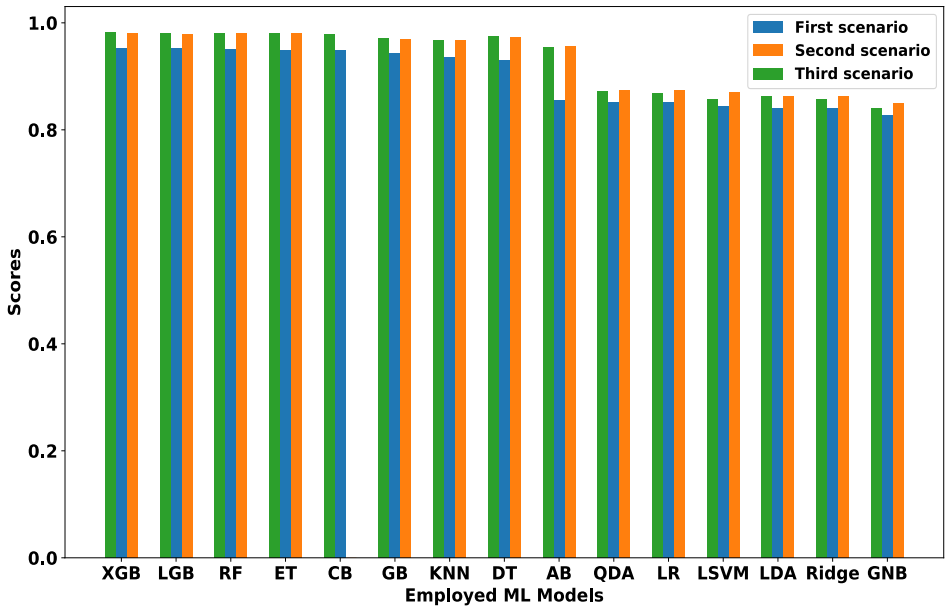
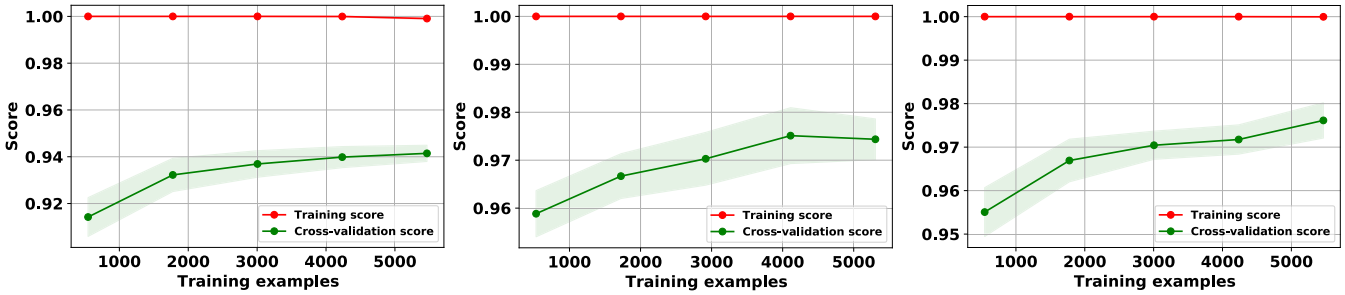


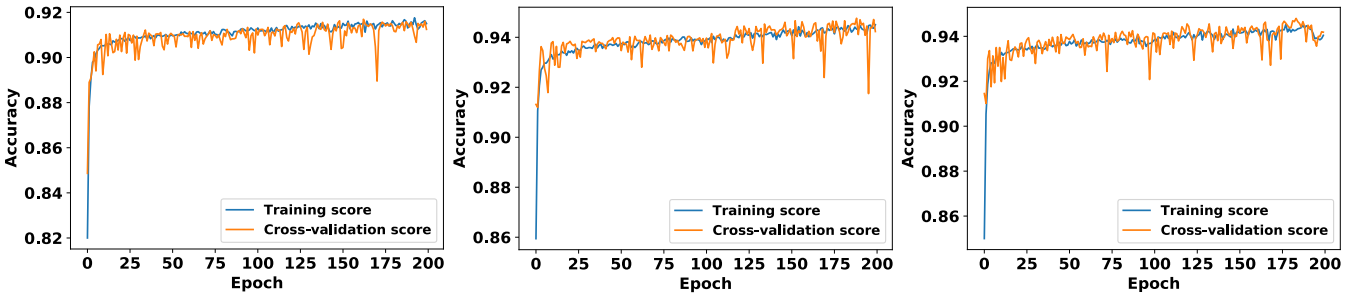
FIGURE 5. Accuracy comparison of linear and nonlinear ML models based on the three scenarios.

Figure 6 shows the learning curves of the proposed model in the three considered scenarios. The obtained training curves are indicated by red color, while the validation curves are represented by green color. The best performance is achieved by the third scenario, in which the earthquakes' class is handled as one entity regardless of being the event



(a) First scenario (XGB) (b) Second scenario (ET) (c) Third scenario (XGB)

FIGURE 6. Training vs. cross-validation comparison: (a) First scenario using "XGB", (b) Second scenario using "ET", and (c) Third scenario using "XGB".



(a) (b) (c)

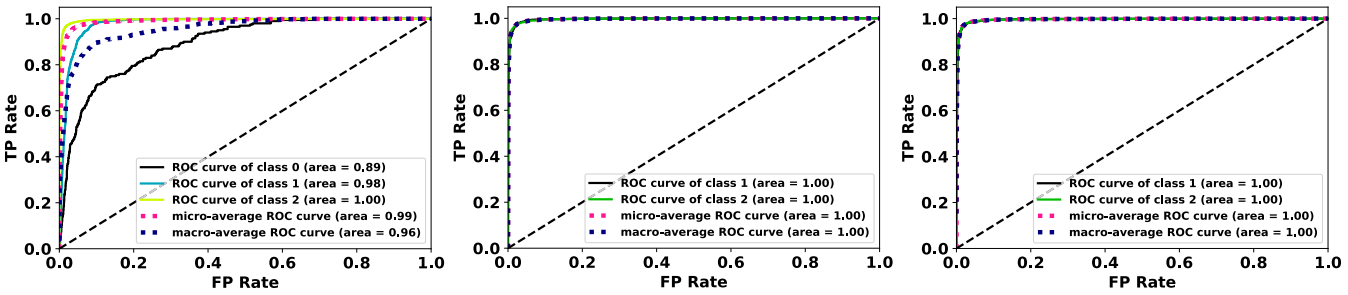
FIGURE 7. Training vs. cross-validation comparison using DL: (a) First scenario, (b) Second scenario, and (c) Third scenario.

is a main-shock or aftershock, and the second class is the artificial seismic sources. This figure is obtained based on the outperforming model in each scenario. More particularly, Figure 6a and 6c are obtained using "XGB" of the first and third scenarios, respectively. Meanwhile, Figure 6b depicts the optimal learning curve of the second scenario obtained by the "ET". Relying on the obtained results, the proposed model can work reliably in the three different considered scenarios. Then, the second and third training performance evaluation comes which is represented by the second and first scenarios, respectively. More particularly, in the least scenario (first) performance, about half of the training period, the curve keeps flat, which means no better performance can be achieved. On the other hand, the performance curves of both the second and the third approximately keep growing, which means these scenarios can reach more enhanced performance. It is clear that there is a gap between the training and validation accuracy. This gap is due to the limited number of the available training samples in the area of interest. We believe that the validation accuracy can be improved if the training samples is increased.

In order to show the performance of DL for solving the considered problem, we introduce the DL results as shown in Figure 7. Specifically, we consider a fully connected neural networks that consists of 6 fully connected layers with 256, 128, 32, 16, 10, and 3 neurons, respectively. The first 5 layers employ the rectified linear unit (ReLU) activation, while the output layer employs the softmax activation. Furthermore, the categorical cross entropy loss function and the adam optimizer are used in the training phase. The DL model is trained for 200 epochs with a batch size of 16.

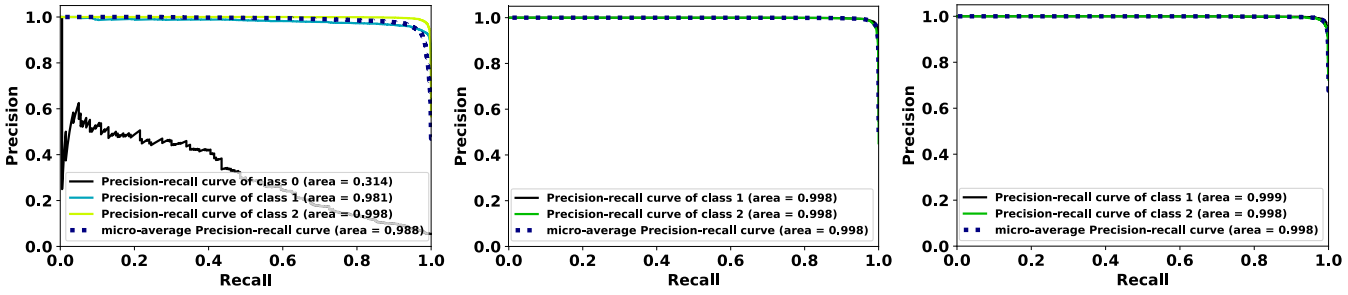
Figure 8 represents the obtained ROC curves of the classes discrimination based on the outperforming model "XGB" in both first and third scenarios, while "ET" in the second one. The ROC curve indicates the trade-off between the true positive rate and the false positive rate of the estimated classes. More concretely, the closer the curve to the left upper corner, the better-class performance is. Throughout the extensive analysis we have done, mostly, the classification performance of the nonlinear models exceeds the corresponding one of the linear models. It is clearly shown that in Figure 8a representing the obtained results of the first scenario, the classification of the artificial seismic sources is the best one by 100% followed by the main-shocks classification with only a 2% decrease. Lastly, the classification of the aftershocks reached 88%. Similarly, Figure 8b and 8c illustrate the obtained sensitivity vs. specificity of the second and third scenarios, respectively. It is obvious that the classification TP rate of the two scenarios is similar, where class "1" and "2" representing earthquakes and artificial seismic sources, respectively, are discriminated by 100%.

Figure 9 indicates the precision vs. recall metric of the three scenarios classification accuracy. It is worth mentioning that after the extensive analysis we only show the precision vs. recall based on the outperforming model among all other ones employed in the classification scheme, which is "XGB" in both the first and third scenarios and "ET" in the second



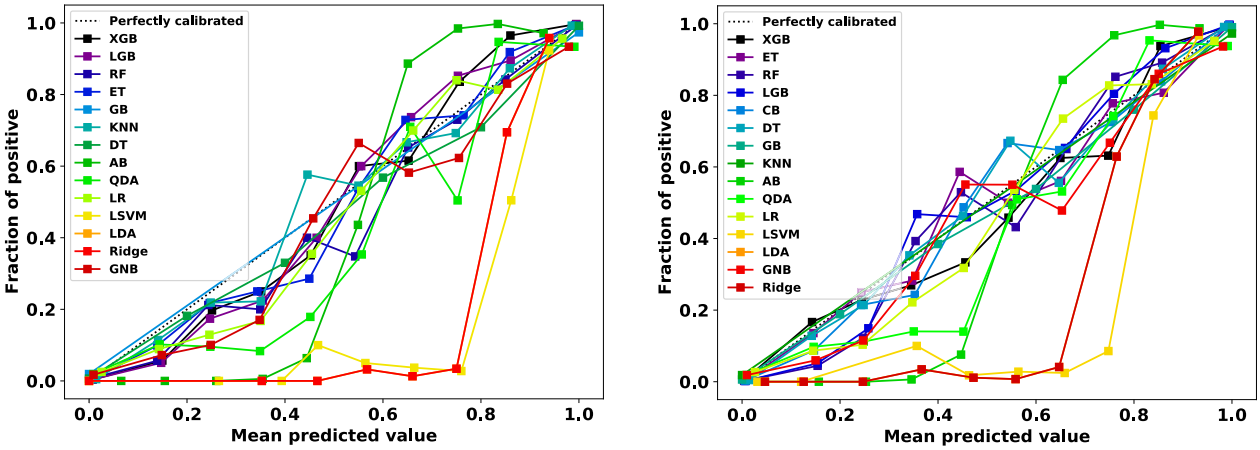
(a) First scenario (XGB) **(b)** Second scenario (ET) **(c)** Third scenario (XGB)

FIGURE 8. ROC curves comparison: (a) First scenario using "XGB", (b) Second scenario using "ET", and (c) Third scenario using "XGB".



(a) First scenario (XGB) **(b)** Second scenario (ET) **(c)** Third scenario (XGB)

FIGURE 9. Precision-Recall comparison: (a) First scenario using "XGB", (b) Second scenario using "ET", and (c) Third scenario using "XGB".



(a) Second scenario **(b)** Third scenario

FIGURE 10. Calibration comparison of linear and nonlinear ML models: (a) Second scenario, and (b) Third scenario.

scenario. First, in Figure 9a, the precision-recall of the artificial seismic sources class (class 2) achieved the optimal value against the main-shocks and aftershocks classes with 98.8%. Then, the main-shocks' class (class 1) is attained with a small decrease as compared to class 2 to be 98.1%, while a steep degradation is obtained with the aftershocks (class 0) by 31.4%. In both second and third scenarios, only two classes are adopted. In other words, the second scenario classes (1 and 2), in which a de-clustering is done, represent the main-shocks and artificial seismic sources, respectively. Meanwhile, in the third scenario no de-clustering is done, where the first class (1) generally represents the earthquakes' class regardless of being that main-shocks or aftershocks, and the second class (2) is the artificial seismic sources. Interestingly, in both Figure 9b and 9c, the precision-recall of artificial seismic sources class reach 99.8%. On the other hand, the precision-recall of the first class (1) in the third scenario is achieved by a little increase as compared to the corresponding in the second scenario with 99.9% and 99.8%, respectively. More particularly, in the third scenario, the de-clustering has not been applied, which means both the main-shocks and aftershocks are all together represented by one class. Accordingly, the available data size for that class is a little larger as compared to the corresponding in the second scenario. Consequently, a slight increase occurred to that class in the third scenario as compared to the second scenario.

Finally, for the two classes' scenarios (second and third), the calibration curves are indicated by Figure 10. The calibration curve reveals the prediction probability, which represents a kind of confidence about the obtained classification results. More particularly, some models converge to poor evaluation of the class probabilities. Accordingly, this calibration curve can reflect the confidence level. Figure 10a indicates the obtained prediction probability confidence level about the case of the second scenario, while the corresponding confidence level of the third scenario is depicted in Figure 10b. In both cases, the use of nonlinear models is beneficial as compared to the linear one. In the second scenario represented by Figure 10a, the ''XGB'' model attains the outperforming confidence, and then the confidence obtained by the ''GNB'' model starts degrading till reaching the minimum confidence. Similarly, the ''XGB'' model presents the best confidence level, while the least confidence is achieved by the ''Ridge'' model as shown in Figure 10b using the third scenario, in which both the main-shock and aftershock are represented by one class and the other is the artificial seismic source.

VI. CONCLUSION

This paper proposed a novel discrimination scheme using ML models based on only four input features (Latitude, Longitude, depth, and magnitude). Throughout this scheme, extensive analyses have been done relying on the common linear and nonlinear ML models. Although, the little number of utilized features, an enhanced classification performance has been attained, which reflects the effectiveness of the proposed scheme. This work also presented an overview of the ML mechanisms in seismicity analysis relying on the seismicity catalog dataset. Moreover, we presented a taxonomy of the linear and nonlinear ML models that can be utilized for such a proposed problem. The obtained results show that the ''XGB'' model outperforms the other employed ML models. The obtained results using DL ensure that DL is not suitable for such designated problem, which relies on only four features; where the proposed ML scheme outperforms the DL. Finally, the proposed scheme is beneficial for efficiently discriminating the contaminated catalogs by the artificial seismic sources, which is not only limited to accurate seismic hazard assessment but also assists in enhancing the planning of future urban developments.

ACKNOWLEDGMENT

The authors would like to acknowledge the support received from Taif University Researchers Supporting Project Number (TURSP-2020/147), Taif university, Taif, Saudi Arabia. They thank the Northern California Earthquake Data Center for making the data is available. They also thank the anonymous reviewers for their valuable recommendations.

REFERENCES

[1] S. Wiemer, ''Mapping and removing quarry blast events from seismicity catalogs,'' *Bull. Seismol. Soc. Amer.*, vol. 90, no. 2, pp. 525–530, Apr. 2000.
[2] J. Murphy and T. Bennett, ''A discrimination analysis of short-period regional seismic data recorded at Tonto Forest Observatory,'' *Bull. Seismol. Soc. Amer.*, vol. 72, no. 4, pp. 1351–1366, 1982.
[3] W.-Y. Kim, V. Aharonian, A. Lerner-Lam, and P. Richards, ''Discrimination of earthquakes and explosions in southern Russia using regional high-frequency three-component data from the IRIS/JSP Caucasus Network,'' *Bull. Seismol. Soc. Amer.*, vol. 87, no. 3, pp. 569–588, 1997.
[4] M. S. Abdalzaher, M. El-Hadidy, H. Gaber, and A. Badawy, ''Seismic hazard maps of Egypt based on spatially smoothed seismicity model and recent seismotectonic models,'' *J. Afr. Earth Sci.*, vol. 170, Oct. 2020, Art. no. 103894.
[5] A. Khodaverdian, H. Zafarani, M. Rahimian, and V. Dehnamaki, ''Seismicity parameters and spatially smoothed seismicity model for iran,'' *Bull. Seismol. Soc. Amer.*, vol. 106, no. 3, pp. 1133–1150, Jun. 2016.
[6] J. Stepp, ''Analysis of completeness of the earthquake sample in the puget sound area and its effect on statistical estimates of earthquake hazard,'' in *Proc. 1st Int. Conf. Microzonazion*, Seattle, WA, USA, vol. 2, 1972, pp. 897–910.
[7] R. E. Habermann, ''Man-made changes of seismicity rates,'' *Bull. Seismol. Soc. Amer.*, vol. 77, no. 1, pp. 141–159, 1987.
[8] H. S. Kuyuk, E. Yildirim, E. Dogan, and G. Horasan, ''An unsupervised learning algorithm: Application to the discrimination of seismic events and quarry blasts in the vicinity of Istanbul,'' *Natural Hazards Earth Syst. Sci.*, vol. 11, no. 1, pp. 93–100, Jan. 2011.
[9] L. Dong, J. Wesseloo, Y. Potvin, and X. Li, ''Discrimination of mine seismic events and blasts using the Fisher classifier, naive Bayesian classifier and logistic regression,'' *Rock Mech. Rock Eng.*, vol. 49, no. 1, pp. 183–211, Jan. 2016.
[10] R. Wang, B. Schmandt, and E. Kiser, ''Seismic discrimination of controlled explosions and earthquakes near mount st. helens using P/S ratios,'' *J. Geophys. Res., Solid Earth*, vol. 125, no. 10, 2020, Art. no. e2020JB020338.
[11] H. S. Kuyuk, E. Yildirim, E. Dogan, and G. Horasan, ''Clustering seismic activities using linear and nonlinear discriminant analysis,'' *J. Earth Sci.*, vol. 25, no. 1, pp. 140–145, Feb. 2014.
[12] D. R. Baumgardt and G. B. Young, ''Regional seismic waveform discriminants and case-based event identification using regional arrays,'' *Bull. Seismol. Soc. Amer.*, vol. 80, no. 6B, pp. 1874–1892, 1990.
[13] G. Horasan, A. B. Güney, A. Käsmezer, F. Bekler, Z. Ötçä, and N. Musaoälu, ''Contamination of seismicity catalogs by quarry blasts: An example from Istanbul and its vicinity, northwestern turkey,'' *J. Asian Earth Sci.*, vol. 34, no. 1, pp. 90–99, Jan. 2009.
[14] M. A. Hedlin, J. B. Minster, and J. A. Orcutt, ''An automatic means to discriminate between earthquakes and quarry blasts,'' *Bull. Seismol. Soc. Amer.*, vol. 80, no. 6B, pp. 2143–2160, 1990.
[15] G. Ataeva, Y. Gitterman, and A. Shapira, ''The ratio between corner frequencies of source spectra of P- and S-waves—A new discriminant between earthquakes and quarry blasts,'' *J. Seismol.*, vol. 21, no. 1, pp. 209–220, Jan. 2017.
[16] A. Yälmaz, Y. Bayrak, and H. Çnar, ''Discrimination of earthquakes and quarry blasts in the eastern black sea region of turkey,'' *J. Seismol.*, vol. 17, no. 2, pp. 721–734, Apr. 2013.
[17] S. Kim, K. Lee, and K. You, ''Seismic discrimination between earthquakes and explosions using support vector machine,'' *Sensors*, vol. 20, no. 7, p. 1879, Mar. 2020.
[18] (2020). *Northern California Earthquake Data Center*. Accessed: Sep. 27, 2020. [Online]. Available: https:/https://ncedc.org/
[19] S. M. Mousavi, S. P. Horton, C. A. Langston, and B. Samei, ''Seismic features and automatic discrimination of deep and shallow induced-microearthquakes using neural network and logistic regression,'' *Geophys. J. Int.*, vol. 207, no. 1, pp. 29–46, Oct. 2016.
[20] N. Rabin, Y. Bregman, O. Lindenbaum, Y. Ben-Horin, and A. Averbuch, ''Earthquake-explosion discrimination using diffusion maps,'' *Geophys. J. Int.*, vol. 207, no. 3, pp. 1484–1492, Dec. 2016.
[21] L. Linville, K. Pankow, and T. Draelos, ''Deep learning models augment analyst decisions for event discrimination,'' *Geophys. Res. Lett.*, vol. 46, no. 7, pp. 3643–3651, Apr. 2019.
[22] X. Shang, X. Li, A. Morales-Esteban, and G. Chen, ''Improving microseismic event and quarry blast classification using artificial neural networks based on principal component analysis,'' *Soil Dyn. Earthq. Eng.*, vol. 99, pp. 142–149, Aug. 2017.
[23] S. S. R. Moustafa, M. S. Abdalzaher, M. H. Yassien, T. Wang, M. Elwekeil, and H. E. A. Hafiez, ''Development of an optimized regression model to predict blast-driven ground vibrations,'' *IEEE Access*, vol. 9, pp. 31826–31841, 2021.
[24] J. Kortström, M. Uski, and T. Tiira, ''Automatic classification of seismic events within a regional seismograph network,'' *Comput. Geosci.*, vol. 87, pp. 22–30, Feb. 2016.
[25] D. G. Kleinbaum, K. Dietz, M. Gail, M. Klein, and M. Klein, *Logistic Regression*. New York, NY, USA: Springer, 2002.

[26] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An Introduction to Statistical Learning*. New York, NY, USA: Springer, 2013, vol. 112.
[27] Y.-W. Chang and C.-J. Lin, "Feature ranking using linear SVM," in *Proc. Causation Predict. Challenge*, 2008, pp. 53–64.
[28] E. Alpaydin, *Introduction to Machine Learning*. Cambridge, MA, USA: MIT Press, 2020.
[29] J. Staal, M. D. Abramoff, M. Niemeijer, M. A. Viergever, and B. van Ginneken, "Ridge-based vessel segmentation in color images of the retina," *IEEE Trans. Med. Imag.*, vol. 23, no. 4, pp. 501–509, Apr. 2004.
[30] H. Chauhan, V. Kumar, S. Pundir, and E. S. Pilli, "A comparative study of classification techniques for intrusion detection," in *Proc. Int. Symp. Comput. Bus. Intell.*, Aug. 2013, pp. 40–43.
[31] A. Pérez, P. Larrañaga, and I. Inza, "Supervised classification with conditional Gaussian networks: Increasing the structure complexity from naive Bayes," *Int. J. Approx. Reasoning*, vol. 43, no. 1, pp. 1–25, Sep. 2006.
[32] T. Hastie, S. Rosset, J. Zhu, and H. Zou, "Multi-class AdaBoost," *Statist. Interface*, vol. 2, no. 3, pp. 349–360, 2009.
[33] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. New York, NY, USA: Springer, 2009.
[34] A. A. Taha and S. J. Malebary, "An intelligent approach to credit card fraud detection using an optimized light gradient boosting machine," *IEEE Access*, vol. 8, pp. 25579–25587, 2020.
[35] T. Chen, "Xgboost: Extreme gradient boosting," *R Package Version 0.4-2*, vol. 1, no. 4, p. 4, 2015.
[36] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.
[37] P. Geurts, D. Ernst, and L. Wehenkel, "Extremely randomized trees," *Mach. Learn.*, vol. 63, no. 1, pp. 3–42, 2006.
[38] M. Xu, P. Watanachaturaporn, P. Varshney, and M. Arora, "Decision tree regression for soft classification of remote sensing data," *Remote Sens. Environ.*, vol. 97, no. 3, pp. 322–336, Aug. 2005.
[39] M. W. Ahmad, M. Mourshed, and Y. Rezgui, "Trees vs neurons: Comparison between random forest and ANN for high-resolution prediction of building energy consumption," *Energy Buildings*, vol. 147, pp. 77–89, Jul. 2017.
[40] S. Tan, "An effective refinement strategy for KNN text classifier," *Expert Syst. Appl.*, vol. 30, no. 2, pp. 290–298, Feb. 2006.
[41] R. Punmiya and S. Choe, "Energy theft detection using gradient boosting theft detector with feature engineering-based preprocessing," *IEEE Trans. Smart Grid*, vol. 10, no. 2, pp. 2326–2329, Mar. 2019.
[42] M. L. McHugh, "Interrater reliability: The kappa statistic," in *Proc. Biochemia Medica*, pp. 276–282, 2012.
[43] J. N. Mandrekar, "Receiver operating characteristic curve in diagnostic test assessment," *J. Thoracic Oncol.*, vol. 5, no. 9, pp. 1315–1316, Sep. 2010.
[44] F. Waldhauser and D. P. Schaff, "Large-scale relocation of two decades of northern california seismicity using cross-correlation and double-difference methods," *J. Geophys. Res., Solid Earth*, vol. 113, no. B8, Aug. 2008.

MOHAMED S. ABDALZAHER (Member, IEEE) received the B.Sc. degree (Hons.) in electronics and communications engineering, in 2008, the M.Sc. degree in electronics and communications engineering from Ain Shams University, Cairo, Egypt, in 2012, and the Ph.D. degree from the Electronics and Communications Engineering Department, Egypt-Japan University of Science and Technology, Madinet Borg Al Arab, Egypt, in 2016.

From 2015 to 2016, he was a Special Research Student with Kyushu University, Fukuoka, Japan. From April 2019 to October 2019, he was with the Center for Japan-Egypt Cooperation in Science and Technology, Kyushu University, where he was a Postdoctoral Researcher. He is currently an Assistant Professor with the National Research Institute of Astronomy and Geophysics, Cairo. His research interests include earthquake engineering, data communication networks, wireless communications, WSNs security, the IoT, and deep learning.

Dr. Abdalzaher is a TPC Member with the Vehicular Technology Conference and the International Japan-Africa Conference on Electronics, Communications and Computers. He is a Reviewer of the IEEE Internet of Things Journal, IEEE Systems Journal, IEEE Access, *Transactions on Emerging Telecommunications Technologies*, *Applied Soft Computing*, the *Journal of Ambient Intelligence and Humanized Computing*, and IET Journals.

SAYED S. R. MOUSTAFA received the B.Sc. degree in geophysics from Cairo University, Egypt, in 1990, the M.Sc. degree in geophysics from Ain Shams University, Cairo, Egypt, in 1997, the Diploma degree in seismology and earthquake engineering from the International Institute of Seismology and Earthquake Engineering (IISEE), Japan, in 2001, and the Ph.D. degree in geophysics from Ain Shams University, Cairo, Egypt, in 2002. Since December 1994, he has been with the Egyptian National Seismic Network Laboratory (ENSN), Department of Seismology, National Research Institute of Astronomy and Geophysics, Cairo, where he was an Assistant Professor, an Associate Professor, in 1998, and a Professor, in 2002. From 2009 to 2019, he was a Professor with the Geology and Geophysics Department, College of Science, King Saud University, Saudi Arabia. His current research interests include earthquake rupture mechanics, numerical methods for wave propagation, spectral element method for ground motion simulation, site response and seismic hazard, characterization of sedimentary basins, and simulation of their seismic response. He is a member of the American Geophysical Union (AGU), the Society of Exploration Geophysicists (SEG), and the Egyptian Geophysical Society (EGS).

MOHAMMED ABD-ELNABY received the B.S., M.S., and Ph.D. degrees in electronic engineering from Menoufia University, Menouf, Egypt, in 2000, 2004, and 2010, respectively. In 2018, he joined the Department of Computer Engineering, College of Computers and Information Technology, Taif University, Al-Hawiya, Saudi Arabia. Since 2010, he has been a Teaching Staff Member with the Department of Electronics and Electrical Communications, Faculty of Electronic Engineering, Menoufia University. Since 2015, he has been working as an Associate Professor with the Department of Electronics and Electrical Communication, Faculty of Electronic Engineering, Menoufia University. He has authored over 70 publications. His research interests include wireless resource management, MAC protocols, cognitive radio, cooperative communication, the IoT, 5G communication, NOMA, and D2D communication.

MOHAMED ELWEKEIL received the B.Sc. degree in electronics and electrical communications engineering from the Faculty of Electronic Engineering, Menoufia University, Egypt, in 2007, and the M.Sc. and Ph.D. degrees in electronics and communications engineering from the Egypt–Japan University of Science and Technology (E-JUST), Alexandria, Egypt, in 2013 and 2016, respectively. He was a Teaching Assistant with the Department of Electronics and Electrical Communications Engineering, Faculty of Electronic Engineering, Menoufia University, from December 2007 to November 2016. In December 2016, he has been promoted to a Lecturer (Assistant Professor) with the Department of Electronics and Electrical Communications Engineering. From April 2014 to March 2015, he was a Research Intern with Alcatel-Lucent Bell N.V. (now NOKIA), Antwerp, Belgium, where he was working on WiFi optimization project. In October 2015, he joined Kyushu University, Fukuoka, Japan, as a Special Research Student for a period of nine months. From April 2018 to March 2020, he was with the College of Information Engineering, Shenzhen University, Shenzhen, China, where he was working as a Postdoctoral Researcher. His research interests include radio resource management for wireless networks, spatial modulation, signal processing for communications, and earthquake engineering.