*Article*

# A Machine Learning Framework for Assessing Seismic Hazard Safety of Reinforced Concrete Buildings

**Ehsan Harirchian** [1],*, **Vandana Kumari** [1], **Kirti Jadhav** [1], **Rohan Raj Das** [1], **Shahla Rasulzade** [2] and **Tom Lahmer** [1]

[1]  Institute of Structural Mechanics (ISM), Bauhaus-Universität Weimar, 99423 Weimar, Germany; vandana.kumari@uni-weimar.de (V.K.); kirti.balasaheb.jadhav@uni-weimar.de (K.J.); rohan.raj.das@uni-weimar.de (R.R.D.); tom.lahmer@uni-weimar.de (T.L.)

[2]  Research Group Theoretical Computer Science/Formal Methods, School of Electrical Engineering and Computer Science, Universität Kassel, Wilhelmshöher Allee 73, 34131 Kassel, Germany; shahla.rasulzade@uni-kassel.de

*  Correspondence: ehsan.harirchian@uni-weimar.de

**Abstract:** Although averting a seismic disturbance and its physical, social, and economic disruption is practically impossible, using the advancements in computational science and numerical modeling shall equip humanity to predict its severity, understand the outcomes, and equip for post-disaster management. Many buildings exist amidst the developed metropolitan areas, which are senile and still in service. These buildings were also designed before establishing national seismic codes or without the introduction of construction regulations. In that case, risk reduction is significant for developing alternatives and designing suitable models to enhance the existing structure's performance. Such models will be able to classify risks and casualties related to possible earthquakes through emergency preparation. Thus, it is crucial to recognize structures that are susceptible to earthquake vibrations and need to be prioritized for retrofitting. However, each building's behavior under seismic actions cannot be studied through performing structural analysis, as it might be unrealistic because of the rigorous computations, long period, and substantial expenditure. Therefore, it calls for a simple, reliable, and accurate process known as Rapid Visual Screening (RVS), which serves as a primary screening platform, including an optimum number of seismic parameters and predetermined performance damage conditions for structures. In this study, the damage classification technique was studied, and the efficacy of the Machine Learning (ML) method in damage prediction via a Support Vector Machine (SVM) model was explored. The ML model is trained and tested separately on damage data from four different earthquakes, namely Ecuador, Haiti, Nepal, and South Korea. Each dataset consists of varying numbers of input data and eight performance modifiers. Based on the study and the results, the ML model using SVM classifies the given input data into the belonging classes and accomplishes the performance on hazard safety evaluation of buildings.

**Keywords:** damaged buildings; earthquake safety assessment; soft computing techniques; rapid visual screening; Machine Learning; vulnerability assessment

## 1. Introduction

A structure's seismic vulnerability is a quantity correlated with its failure in the event of earthquakes of previously known intensity. This quantity and seismic hazard experience's importance helps us determine the potential risk from future earthquakes [1]. Many destructive seismic activities

have caused immense disruption in human history, resulting in substantial loss of life, severe economic consequences, and severe property damage. There has been a remarkable increase in working personnel's migration from rural to urban metropolitan regions due to career prospects and lifestyle. As a consequence, this imparts the responsibility of protecting the high occupancy of urban infrastructure. Old buildings still in service, historical structures with heritage, high importance buildings, and buildings not compliant with the latest seismic codes are the buildings with the highest seismic vulnerability. This proves the necessity of manifesting a seismic structural prioritization scheme, which shall prevent damage or adapt to post-disaster management regulations. Rapid Visual Screening (RVS) is a rapid and reliable method for determining the damage index for various buildings [2,3]. Sinha and Goyal [4] have a laconic discussion as motivation for a layman. The United States first recognized the demand for a fast, reliable, and computationally easy method; therefore, the first RVS method was mentioned in 1988 as "Rapid Visual Screening of Buildings for Potential Seismic Hazards: A Handbook" [5]. This RVS approach (1988) was subsequently updated in 2002 with new developments in seismic engineering. However, an in-depth study and detailed analysis bring challenges in the urban sector where buildings are complex and built on the absence of non-standard norms of safety. The research focuses on one of the latest techniques for enhanced vulnerability assessment of reinforced structures. Eventually, many other countries followed this strategy by changing their RVS approaches concerning their local circumstances, improvements, etc. [6]; for instance, Indian RVS (IITK-GSDMA) [7] or the Philippine RVS [8].

The approach of RVS usually begins with a walk down evaluation to perform a visual inspection physically for recording the seismic parameters. For this purpose, the collection of datasheets needs an evaluator. RVS is a highly effective mechanism to prioritize the structures with high vulnerability due to the faster computational time and non-tedious methodology, thereby helping analyze more massive building stock in a short time. The RVS method is based on a scoring system, wherein the final performance score shall be obtained by performing elementary computations. Every singular approach has the respective cut-off score predefined, for instance, the Federal Emergency Management Agency (FEMA). The structures that do not reach the cut-off score shall be subjected to comprehensive second and third assessment stages. We shall note that seismic vulnerability assessment is performed mainly in three stages [9]: walk down an evaluation, preliminary evaluation, and detailed evaluation. The first stage is defined by a physical walk down evaluation with fundamental computations for the screening of vulnerable structures. Structures that struggle to fulfill expectations at this stage will be addressed with further action. The second step consists of a comprehensive analysis conducted by the detailed study of several structure components, such as the actual ground conditions, the quality of the material used, the state of the structural elements, etc. The third stage is brought into action when the structures require a more detailed analysis. The concept of structural activity in seismic activities, defined as non-linear dynamic structural research, is performed here [10].

A substantial literature has been published in the attempts of integrating several methods from various domains with RVS, for instance, statistical methods [11,12], Artificial Neural Network (ANN) [13–17], multi-criteria decision making [18,19], and type-1 [20–23] and type-2 [24,25] fuzzy logic systems are frequently assimilated within RVS for increasing the interface and efficacy of seismic vulnerability screening. However, there are methods developed to evaluate the damage and change detection of buildings by remote sensing and image analysis [26,27]. The parameter significance and the selection of the optimum number of parameters have been used effectively by Morfidis and Kostinakis [28]. Multiple linear regression analysis [11] is the most commonly used statistical technique for the classification of damage state in the RVS domain, preceded by other approaches, such as discriminant analysis suggested in [29]. Tesfamariam and Liu [15] have effectively applied and utilized eight different statistical risk classification methods, including six-building output modifications (predictor variables) with damage states as none (O), light (L), moderate (M), severe (S), and collapse (C). Furthermore, Jain et al. have proved the integrated use of various exciting variable selection techniques such as $R^2$ adjusted, forward selection, backward elimination, and Akaike's and Bayesian

information criteria [30]. Besides, the traditional least square regression analysis and multivariable linear regression analysis were used in the research proposed in [31]. Several probabilistic models are suggested, such as reliability-based models and "best estimate" matrices of the potential damage [32].

Morfidis and Kostinakis [33] have proposed a unique and innovative ANN application from a practical perspective. The coupling of fuzzy logic and ANN's has been used in a study by Dristos [34], where untrained fuzzy logic procedures are remarkable. Nevertheless, several of the previous RVS approaches are based on the expert's view, complexities, or conclusions in the linear parameter relationship. Machine Learning (ML), an artificial intelligence division, uses computational algorithms capable of automatic creation by training and learning. This focuses on rendering the forecasts using data analysis software. The ML algorithms generally comprise of three fundamental components: representation, evaluation, and optimization. A method known as Support Vector Machine (SVM) has been commonly used in scenarios of processing uncertain data or data without necessary records that need to be processed efficiently. SVM performs adequately even for the provided information in the context of unstructured and semi-structured data, such as tests of images merged with trees. Kernel trick is one of SVM's strengths that integrates all the knowledge required for the learning algorithm to identify a core object in the transformed field [35]. The implementation of SVM with initially supervised learning data-sets is seen in this analysis. The damage records of 4 different earthquakes have been taken to implement from countries such as Ecuador, Haiti, Nepal, and South Korea. Finally, the work demonstrates the SVM model's expertise and the detailed interpretation of the chosen earthquakes result.

## 2. Background of the Selected Earthquakes

In this study, the archived data of buildings on the open-access database on the Datacenterhub platform [36] were used. These buildings were affected during Nepal, Haiti, South Korea, and Ecuador earthquakes, and they were investigated further by different research groups, and their information and observed damage were collected.

The entire province across Ecuador is prone to seismic hazards, and in this study, one of the current earthquakes, which occurred on April 16, 2016, and majorly affected the coastal zone of Manabí, has been selected. The data collection has been carried through the American Concrete Institute (ACI) research team to generate a memorandum of damaged RC structures affected due to the earthquake. These teams collaborated with the technical staff and students of the Escuela Superior Politécnica del Litoral (ESPOL) [37]. The data have two different soil profiles known as APO1 (the location of the IGN strong-motion station) and Los Tamarindos. APO1 is situated amid alluvial soils and colluvial deposits with an average shear wave velocity at the height of 30 m, Vs30 of 240 m/s, whereas Los Tamarindos is within the alluvial clay and silt deposits with, Vs30 of 220 m/s [38].

Secondly, the Haiti earthquake that occurred on January 12, 2010, has been taken for the assessment. The seismologists have compiled the structural data of 145 RC frames from the University of Purdue, Washington University, and Kansas University have been worked together with research workers from Université d'Etat d'Haïti [39,40]. Soil properties of the data are classified as the granular alluvial fans and the soft clay soils for the southern and the northern plain of Léogâne city, respectively, which is located to the west of Port-au-Prince with the severe structural damage [41]. According to the Shake-Map of the U.S. Geological Survey, Léogâne has a ground motion with the highest instrumental intensity of IX, and Port-au-Prince has intensity VIII [42].

The devastating earthquake which had struck Nepal in May 2015 has also been assessed in the current study. Researchers from the Purdue University and Nagoya Institute of Technology, in association to the ACI, conducted surveys and inspected damaged RC structures. A dataset compilation has been completed in the reconnaissance survey, which has 135 low-rise reinforced concrete buildings with or without masonry infill walls [43]. The fill sediments of the Kathmandu valley's soil profile mostly consist of a heterogeneous mixture of clays, sands, and silts, which have a

thickness of nearly 400 m [44]. Based on the records of the ShakeMap of the U.S. Geological Survey, the ground motion intensity of VIII with the epicenter of 19 KM to the south-east of Kodari.

Ultimately, the earthquake emerged on November 15, 2017, in the Korean territory. Two mega cities Heunghae and Pohang, were influenced by the earthquake event, and due to that, structures were considerably damaged in the vicinity of these areas. Additionally, thousands of inhabitants were homeless caused by severe structural damage, leading to a significant impact on the economy of around 100 million dollars in public and private infrastructures. The collection of damage data had been carried by a team of researchers from ACI in collaboration with multiple universities and research institutes [45].

The subsurface soil strata consist of filling, alluvial soil, weathered soil, weathered rock, and bedrock. The ShakeMap of the U.S. Geological Survey has listed the earthquake intensity as VII [46]. Table 1 indicates the details of selected earthquakes and related parameters.

**Table 1.** Properties of the selected earthquakes.

| Earthquake | Year | Moment Magnitude ($M_w$) | PGA (g) | PGV (cm/s) |
|---|---|---|---|---|
| Ecuador [47] | 2016 | 7.8 | 0.4 | 82 |
| Haiti [39] | 2010 | 7 | 0.4 | 40 |
| Nepal [48] | 2015 | 7.3 | 0.2 | 20 |
| S. Korea [49] | 2017 | 5.4 | 0.2 | 10 |

*Choice of Building's Damage Inducing Parameters*

It is necessary to gather certain primary information beforehand, to begin with, every RVS process. Several experiments have investigated the usefulness of building characteristics as input variables for evaluating seismic susceptibility [18,50,51]. In line with FEMA154 [5], many of the methods as mentioned earlier contained parameters and exhibited that the most useful inputs are namely, (i) system type, (ii) vertical irregularity, (iii) plan irregularity, (iv) year of construction, and (v) quality of construction. Yakut et al. [52] considered more criteria for the seismic risk assessment of buildings. Concerning the characteristics of the affected structures and the tremendous amount of current building stocks, the following eight criteria were formulated for research purposes. The criteria considered are No. of story, total floor area, column area, concrete wall area in X and Y directions, masonry wall area in X and Y directions, and captive columns. These parameters also form the basis of the Priority Index (PI), as proposed by Hasan and Sozen [53]. These parameters have been selected because the effect of these has been examined and adjusted to assess structures, not meeting code requirements. The parameters can also be applied to different regions as they have been tested in areas with different construction practices. Furthermore, the parameters are easily collected by visual inspection, which leads to a lowered vagueness for investigators, thereby saving time. The eight parameters have been presented in Table 2.

**Table 2.** Parameters for earthquake hazard safety assessment [16].

| Variable | Parameter | Unit | Type |
|---|---|---|---|
| $\alpha_1$ | No. of story | N | Quantitative |
| $\alpha_2$ | Total Floor Area | $m^2$ | Quantitative |
| $\alpha_3$ | Column Area | $m^2$ | Quantitative |
| $\alpha_4$ | Concrete Wall Area (Y) | $m^2$ | Quantitative |
| $\alpha_5$ | Concrete Wall Area (X) | $m^2$ | Quantitative |
| $\alpha_6$ | Masonry Wall Area (Y) | $m^2$ | Quantitative |
| $\alpha_7$ | Masonry Wall Area (X) | $m^2$ | Quantitative |
| $\alpha_8$ | Captive Columns | $N$ (exist = 1, absent = 0) | Dummy |

## 3. ML Approach for Multi-Class Classification

ML classifies objects quite fast and efficiently, which humans may find it difficult to analyze because of complex numbers, vast data, or a huge number of features. ML algorithms are primarily for binary classifications. However, when added to the present algorithms, several techniques allow successfully to segregate the multi-class entity. SVM strongly solves multi-class classification or regression problems for wide assorted fields, like Anomaly Detection, text or image categorization, time-series analysis, and medical informatics [54]. The study has enforced the SVM algorithm on the datasets using techniques like One-vs.-All [55] and GridSearchCV to build an SVM classifier for multi-classes classification efficiently.

Figure 1 presents an approximate procedural used in the study. The workflow comprises data pre-processing, data splitting, model building, model fitting, accuracy optimization, and visualization; each step is elaborated in the following sub-sections.
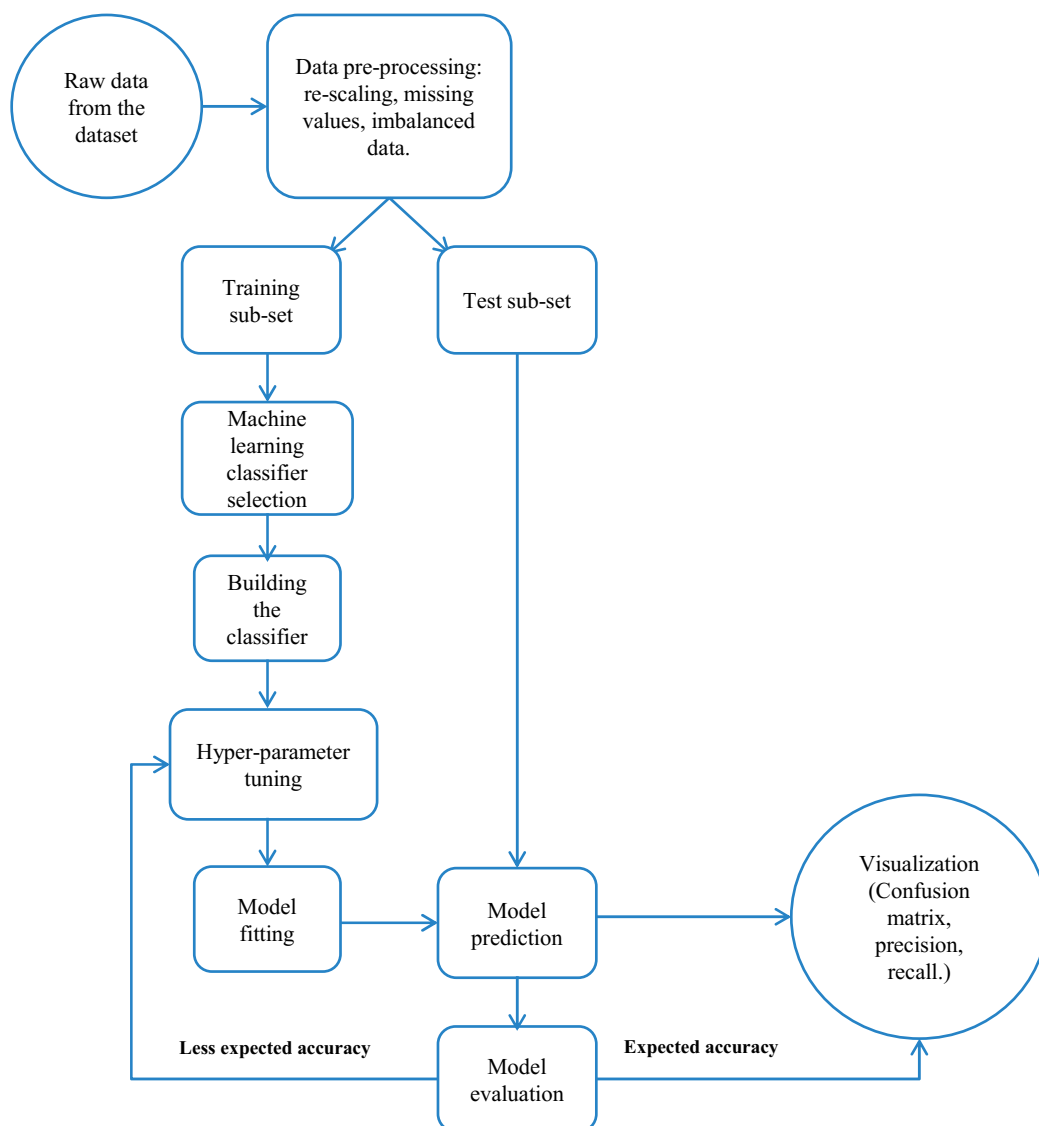


**Figure 1.** Flow chart for Machine Learning (ML) methodology implementation.

## 3.1. Data Pre-Processing

Data are an integral part of any ML classifier, and they are composed of input features as independent variables and labeled output as the dependant variable. The quality of data and the information derived from a dataset influence the learning capacity of the model. Therefore, pre-processing before feeding the data to the model is essential. Row data may contain null values, outliers, and non-categorical data. ML is taught to classify only using categorical data, and it can not process ordinal or numeric datatype. Library *"Panda"* is well-known in such a situation for data manipulation and analysis. It fits 1 for the input observations that belong to the particular class, leaving the rest of the observation 0. Similar function is performed by *One-Hot-Encoding* in *sklearn* library. ML treats missing value as *NaN* also known as *not-a-number*. There are several ways to deal with missing data, like omitting the complete row to which the missing value belongs, or replacing the missing value with mean, median, or most occurred observation. *SimpleImputer* from *sklearn* easily imputes the missing data with a constant value or using statistics like mean, median, or most common value. Often the dataset accommodates attributes with a mixture of scale. ML models anticipate or are more efficient if the data follow the same scale. Data standardization or normalization are commonly practiced methods for data re-scaling, availed in *sklearn*. Data standardization is the distribution of attribute between a mean of zero and a standard deviation of one (unit variance), whereas data normalization is scaling of attributes into the range 0 and 1.

## 3.2. Imbalanced Data

A frequent problem in ML classification is a disproportionate ratio among the number of attributes in each class, called as imbalanced classes [56]. The classifier's efficiency and accuracy are optimal when an equal number of samples are present for each class. A classification problem may be slightly skewed if there is a small imbalance. On the other hand, classification problems could face significant disagreement if the number of examples is enormous for one class and very few in another class for a given training dataset. Random sampling is a common approach used to address the problem of imbalanced data. Several proven approaches help deal with imbalanced data. Like, Random Under-sampling is a random deletion of excess examples in the majority class, and Random Over-sampling is random duplication of examples in the minority class. Random Under-sampling is not a preferred method when the data are crucial, and discarding or deleting data can cause information loss. *SMOTE*, i.e., *Synthetic Minority Over-sampling Technique* from library *imblearn* works well in most scenarios as it merely adds examples in minority class without adding any new information to the class.

## 3.3. Classification

Machine Learning is well-known for classification problems, either binary or multi-class classification. Depending on the case study, the selection of the input parameter varies. In general, including a few essential parameters create a better impact in accuracy than a high number of less critical input parameters.

The classification of the buildings' damage states is ranked to anticipate the risk associated with each structure. Table 3 summarizes the representation of damage states and their associated damage used in the study.

**Table 3.** Damage classification (adopted from [57]).

| Damage Grade | Damage State | Damage Visibility |
|---|---|---|
| 1 | None | No damage |
| 2 | Light | Cause hairline cracks in the structure; resigned plasters |
| 3 | Moderate | Breaking plasters; teared walls and visible joints between panels |
| 4 | Severe | Reduced regional structural and wide wall plans |

*3.4. Splitting of Dataset*

Splitting the dataset appears as essential to overcome the bias towards training data in ML algorithms. ML classifiers often overfit the data to fit best the training data, which results in poor prediction performance on actual test data. The training subset creates the predictive model, and the test subset evaluates the performance of the model classifier. As a good practice, 80% of the data goes to the training subset, and 20% falls towards the test subset.

*3.5. SVM as Classifier*

ML algorithms are broadly classified among supervised learning and unsupervised learning. SVM is a popular supervised learning algorithm and is frequently used to evaluate data for classification and regression analysis. SVM is primarily designed only for two-class distribution [58], but presently the algorithm is advanced and successfully works with multi-class classification problems using *kernel tricks* [59] and *soft margin*. This study involves kernel trick with hyperparameter tuning (explained further) for classifying the multi-class building samples to their respective classes. Figure 2 shows the elementary SVM classifier mechanism while evaluating two linearly separable classes; for instance, the feature points are considered belonging to the positive class and negative class. SVM takes the feature points and draws the hyperplanes, which separates the points into their belonging classes. An optimal hyperplane or decision boundary is the one that fits best to segregate the feature points. Margin determines the distance of the decision boundary from the nearest data points. Support vectors are the feature points closest to the optimum hyperplane, located at a minimum distance from it.

The hyperplane separating the classes can be represented as [60]:

$$W_0{}^T x + b_0 = y \tag{1}$$

where, W is the *weight* to be learn using the training samples,
$x$ denotes the feature vector; $x \in \mathbb{R}^n$,
$b$ is a constant,
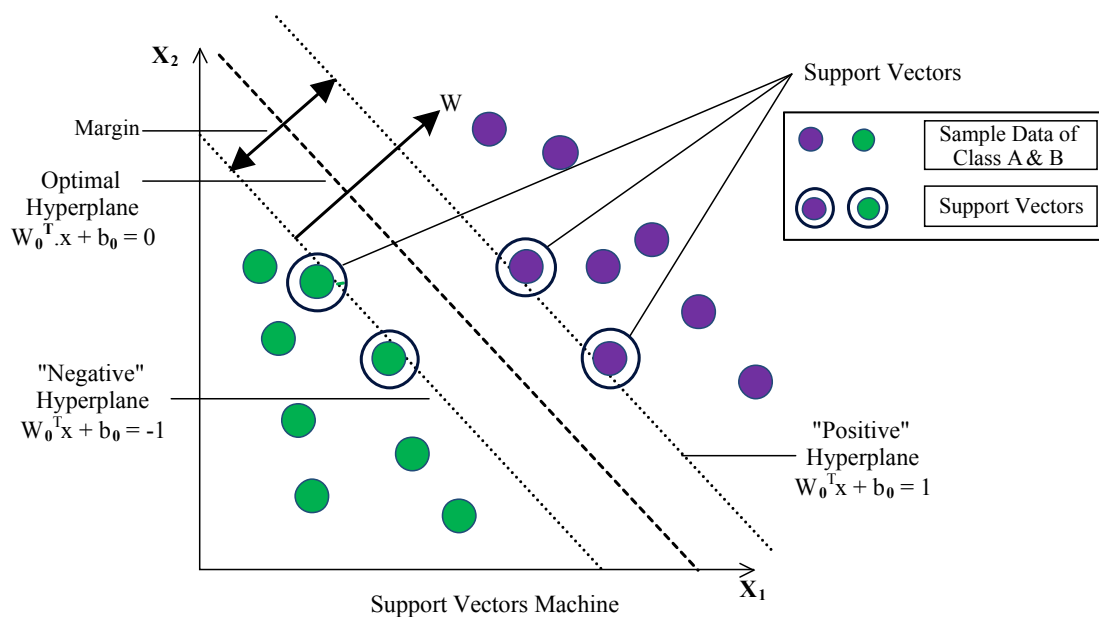$y$ is the class value of the training samples; $y \in \{1, -1\}$.



**Figure 2.** Support Vector Machine: Mechanism.

*3.6. Performance Evaluation and Utilization of Model*

Performance evaluation is an intrinsic part of the classifier modeling process. It finds the optimal classifier representing the sample data and how precisely the built model will perform future data. Performance evaluation of any ML classifier aims to anticipate the prediction accuracy on future or unseen samples. Evaluating the model accuracy with the training set data can overfit the model.

For proper performance interpretation of a model, some part of the dataset is kept on hold as test subset, although the test subset has labeled output. The model gets trained on the training sub-set, and further, the held-out observations are sent to the model to predict. Known labels of the test sub-set are then compared with the prediction returned by the model.

## 4. Methodology implementation

The study includes data from four different countries, including Ecuador, Haiti, Nepal, and South Korea, with varying numbers of RC building samples, as mentioned earlier from Datacenter hub platform [36,37,40,43,45]. Table 4 presents the distribution of sample data, featured parameter, and damage classes in the given dataset. Each dataset contains eight parameters: number of stories, total floor area, column area, concrete wall area in X and Y direction, masonry wall area in X and Y direction, and captive columns. The damage classes show the severity of the damage caused post the earthquakes.

The datasets show an uneven distribution of samples for each damage class. Seismic susceptibility of similar buildings is forecasted based on the relevance damage class. Ecuador and Haiti samples are classified among three different damage classes, whereas Nepal and South Korea samples have four damage classes. The distribution of samples among the damage scales for each dataset is shown in Figure 3.
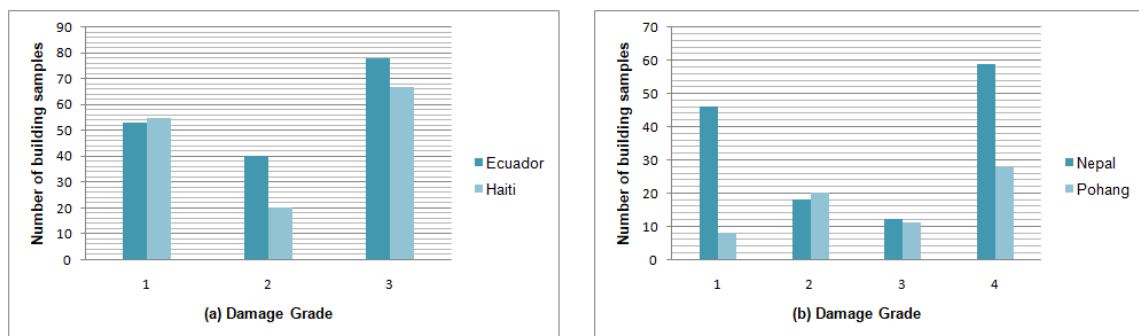


**Figure 3.** (**a**) Data for Ecuador and Haiti and (**b**) the data for Nepal and South Korea. The distribution of sample buildings over the damage classes is non-uniform. South Korea has the least number of samples among the other datasets.

**Table 4.** Data distribution in the dataset.

| Country | Building Samples | Feature Parameters | Damage Classes |
|---|---|---|---|
| Ecuador | 171 | 8 | 3 |
| Haiti | 142 | 8 | 3 |
| Nepal | 135 | 8 | 4 |
| South Korea | 67 | 8 | 4 |

With the histogram's help, the distribution of data of all the eight parameters within a dataset is assessed. Figure 4 represents the histogram for Ecuador. The abscissa of the plots represents the scale between the highest and the lowest value of each parameter. The frequency distribution of the values shows non-linearity in the variation; rather, column area, Masonry Wall Area (NS), number

of floors, and total floor area show Gaussian distribution yet skewed to the left. Masonry wall Area (EW) is distinct from an exponential distribution. Concrete wall area for Y and X directions show no data, while captive columns are distributed bimodally. Haiti, Nepal, and South Korea follow similar non-linear behavior for data distribution.
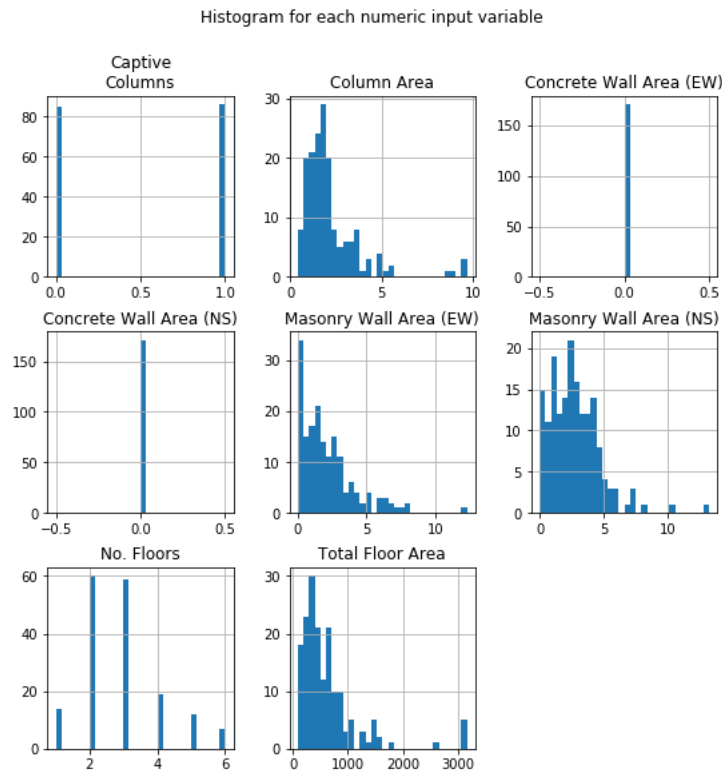


**Figure 4.** Ecuador's histogram plots show that each parameter's data or value in the input dataset spreads disorderly. The result affects the stability and predictability of the classifier. The other three input datasets follow the identical pattern.

*4.1. Data Pre-Processing*

Data pre-processing is critical while feeding the data to the model for learning. There are various ways to manipulate the data as per the requirement; a few important ones are considered for shaping the datasets used in the study.

- Re-scaling the data: To accommodate the data on the same scale, the data are standardized using *"MinMaxScaler"* from *sklearn* library. The standardization frames all the input features between 0 and 1, which brings robustness to minimal standard deviation of features and speeds up the algorithm's calculation.
- Strategy for missing data: The datasets contain very few missing attributes. Removal of data could cause the study of a critical loss of information like the seismic vulnerability test of buildings. Therefore, the missing values are replaced by *mean* strategy using *SimpleImputer* from *sci-kit learn* library.
- Imbalanced data: The imbalanced data in the dataset are handled by Random Over-sampling using *SMOTE*. It randomly duplicates the minority class's attributes to match the majority class's sample number, but do not make any effective change in the class. Tables 5–8 shows how the imbalanced data get manipulate with each iteration for the dataset of Ecuador, Haiti, Nepal, and South Korea, respectively. The final iteration in all datasets has an equal number of data in each class.

**Table 5.** Ecuador—Random Over-sampling of data [61].

| Iteration | Class I | Class II | Class III |
|---|---|---|---|
| 0 | 53 | 40 | 78 |
| 1 | 53 | 78 | 78 |
| 2 | 78 | 78 | 78 |

**Table 6.** Haiti—Random Over-sampling of data [39].

| Iteration | Class I | Class II | Class III |
|---|---|---|---|
| 0 | 55 | 20 | 67 |
| 1 | 55 | 67 | 67 |
| 2 | 67 | 67 | 67 |

**Table 7.** Nepal—Random Over-sampling of data [62].

| Iteration | Class I | Class II | Class III | Class IV |
|---|---|---|---|---|
| 0 | 46 | 18 | 12 | 59 |
| 1 | 46 | 18 | 59 | 59 |
| 2 | 46 | 59 | 59 | 59 |
| 3 | 59 | 59 | 59 | 59 |

**Table 8.** South Korea—Random Over-sampling of data [63].

| Iteration | Class I | Class II | Class III | Class IV |
|---|---|---|---|---|
| 0 | 8 | 20 | 11 | 28 |
| 1 | 28 | 20 | 11 | 28 |
| 2 | 28 | 20 | 28 | 28 |
| 3 | 28 | 28 | 28 | 28 |

### *4.2. Splitting of Dataset*

For all the four datasets, the training and test sub-sets divide into 80% and 20%, respectively. The training set contains all the attributes with labeled output, whereas the test sub-set evaluates the learning performance.

### *4.3. Hyperparameter Tuning—SVM*

While building an ML classifier, there are design choices that define the architecture of the classifier. These choices or parameters are known as hyperparameters, and to achieve an optimal model; the hyperparameters require tuning. Hyperparameters often manage the learning method and the learning pace of the model. Hyperparameters are different from the model parameters, and they can not get trained directly from the data.

SVM has important hyperparameters like kernels, C, and gamma. Kernels modify the training samples to transform any non-linear function into a higher dimension linear function. Radial Basis Function Kernel or RBF, Sigmoid, Polynomial or poly, and Gaussian are few kernel types that SVM uses. Poly and RBF are better suited for multi-class classification problems. Gamma determines the impact of a new feature point on the decision boundary. The high value of gamma means close, and the low value means far from the decision boundary. C acts as a penalty factor for the classifier. Therefore, every time the classifier misclassifies any sample data, the C value increases. Tuning the hyperparameters and finding the optimal accuracy in the model is not cost-effective if processed intuitively. *Grid search* is reasonably the fundamental tuning approach. Various possible combinations of all hyperparameters are fed into the grid search, and the model architecture with the best accuracy is selected.

## 5. Result and Discussion

The model's certainty to predict the class of the unknown samples correctly defines its efficiency, illustrated by various factors, like confusion matrix, accuracy, recall, and precision. The confusion matrix visualizes the performance of the model by producing numbers of samples predicted correctly and incorrectly. Precision, also known as Positive Predictive Value or PPV, is the fraction of cases when the model identifies the class of the sample correctly. Accuracy is the measure of correctly predicted sample points out of the complete data. Recall or sensitivity is the fraction of cases rightly identified as belonging to the respective class amidst all the samples that correctly fit that class. While evaluating the four datasets, all the four mentioned factors are considered. Figures 5–8 present the confusion matrix and calculated Precision and Recall for each class for datasets of Ecuador, Haiti, Nepal, and South Korea. The ($3 \times 3$) or ($4 \times 4$) size of the confusion matrix follows the number of classes in the respective datasets.

|  | Class 1 | Class 2 | Class 3 | Recall |
|---|---|---|---|---|
| Class 1 | 10 | 4 | 3 | 59% |
| Class 2 | 4 | 11 | 1 | 69% |
| Class 3 | 4 | 2 | 7 | 54% |
| Precision | 55% | 65% | 64% | 60% |

**Figure 5.** The overall accuracy attained by the classifier for Ecuador is 60%. Class 2 has the maximum correct predictions; wheres class 3 happens to miscategorize significant unseen samples.

Haiti model in Figure 6 shows the highest accuracy of 68%, and precision of 58%, 91%, and 64% for damage classes 1, 2, and 3, respectively. Class 1 building samples are efficiently classified, i.e., 11 out of 19, while Class 3 samples show a low acceptance, and only 7 out of 17 are predicted accurately.

|  | Class 1 | Class 2 | Class 3 | Recall |
|---|---|---|---|---|
| Class 1 | 11 | 0 | 0 | 100% |
| Class 2 | 3 | 10 | 4 | 59% |
| Class 3 | 5 | 1 | 7 | 54% |
| Precision | 58% | 91% | 64% | 68% |

**Figure 6.** Haiti classifier scored the highest accuracy of 68%. Class 1 and class 2 samples are predicted correctly; wheres class 3 samples have maximum misclassified predictions. Class 1 has 100% recall value, i.e., no false-positive prediction, and 91% of precision for class 2 shows that the classifier significantly identifies the class samples accurately.

|  | Class 1 | Class 2 | Class 3 | Class 4 | Recall |
|---|---|---|---|---|---|
| Class 1 | 10 | 1 | 1 | 2 | 71% |
| Class 2 | 2 | 8 | 3 | 1 | 57% |
| Class 3 | 0 | 0 | 11 | 0 | 100% |
| Class 4 | 3 | 1 | 2 | 3 | 33% |
| Precision | 67% | 80% | 65% | 50% | 67% |

**Figure 7.** With four damage classes, Nepal scored the second-highest accuracy of 67%. Class 3 has maximum true-positive samples with a recall value of 100%.

Among all the models, the South Korea model functioned with the least accuracy of 48%. A low number of sample data can be one of the reasons for the model's lower accuracy. Class 1 and Class 3 have the highest and same recall value of 67%; it shows that these classes are correctly recognized, while class 4 is poorly misclassified with a recall value of only 20%.

|  | Class 1 | Class 2 | Class 3 | Class 4 | Recall |
|---|---|---|---|---|---|
| Class 1 | 2 | 0 | 1 | 0 | 67% |
| Class 2 | 0 | 2 | 4 | 0 | 33% |
| Class 3 | 0 | 1 | 6 | 2 | 67% |
| Class 4 | 1 | 2 | 1 | 1 | 20% |
| Precision | 67% | 40% | 50% | 33% | 48% |

**Figure 8.** South Korea performed with the least accuracy of 48% among all the other classifiers. Class 1 has a moderate recall and precision value, i.e.; 67%, whereas class 4 acquired the least recall and precision values.

To understand the practicality of balancing the imbalanced data, the study evaluated the model prediction efficiency before and after data manipulation. Table 9 shows the accuracy of each model with the current kernel, achieved prior and later shaping of the imbalanced data. The accuracy of each model improved after balancing the imbalanced data. The highest accuracy of 68% for the Haiti model was only 60% when the data were imbalanced, whereas South Korea only had 43% accuracy, even lesser than 48%. Most results employed RBF kernels.

**Table 9.** Model accuracy with and without imbalanced data.

| Country | Imbalanced Data Accuracy and Kernel (%) | Balanced Data Accuracy and Kernel (%) |
|---|---|---|
| Ecuador | 54 RBF | 60 RBF |
| Haiti | 52 RBF | 68 RBF |
| Nepal | 59 Sigmoid | 67 RBF |
| South Korea | 43 RBF | 48 RBF |

The Receiver Operating Characteristics (ROC) curve is represented by a graphical plot that depicts a binary classifier model's indicative capacity as its bias threshold differs. The *X*-axis and *Y*-axis of the ROC curve display true positive and false positive, respectively. Therefore, the ROC plot's top left corner acts as an "ideal" point—depicting value zero for a false positive rate and a true positive rate. Therefore, the larger the area under the curve (AUC), the better is the classifier performance. The ROC curve's steepness serves an essential part in expanding the true positive rate and reducing the false-positive rate. Since the ROC curve works only in binary classification, a multi-class classification method needs to binarize the output. The curves of different classes generated in ROC are comparable with each other or with a varying threshold.

Figures 9–12 illustrates the ROC plots for the test data for Ecuador, Haiti, Nepal, and South Korea models, respectively. *Macro-average* shown in the ROC curves is to evaluate multi-class classification problems. Macro-average computes the test data independently for each class and calculates the average.

ROC curves for Haiti (Figure 10) and Nepal (Figure 11) illustrate the better performance of their model over the test data with a probability of 85% and 84%, respectively. Figure 10 clearly shows that the model significantly classified test samples belonging to class 2 (95%), whereas test samples belonging to class 3 are moderately predicted (74%). As macro-averaging is the average of model performance for each class, Ecuador shows the least macro-average percentage (76%) with an efficiency of only 53% in classifying test samples belonging to the class 3. The proposed method results show a significant improvement in RVS methods such as RVS based on multi-criteria decision-making [19] or Multi-Layer Perceptron [13] where the accuracies were around 37% and 52%, respectively.
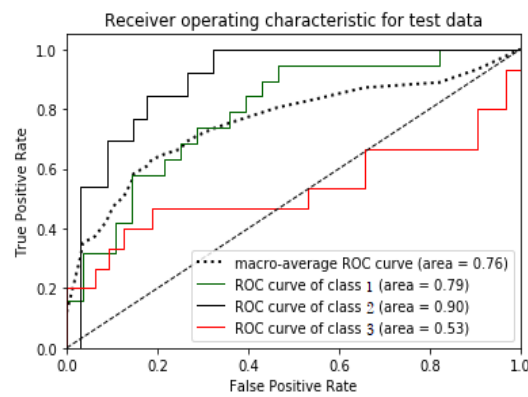
**Figure 9.** Ecuador—ROC curves for each class of test dataset. The sample data for classes 1 and 2 are classified better than class 3, as the curve for class 3 is below the threshold. The result also indicates that the false positive rate for class 3 is high.
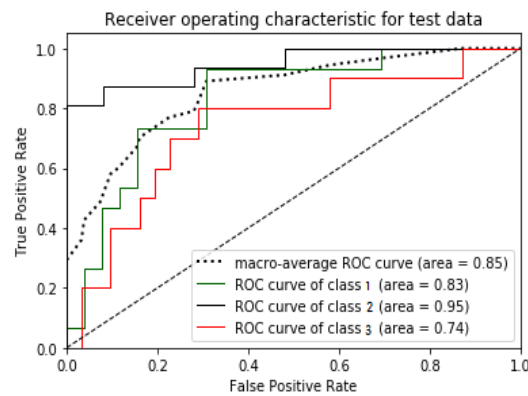


**Figure 10.** Haiti—ROC curves for each class of test dataset. Haiti's classifier has performed well in classifying the sample data for each class except class 3, as all the curves are above the threshold. There are few false-positive appearing for class 3.
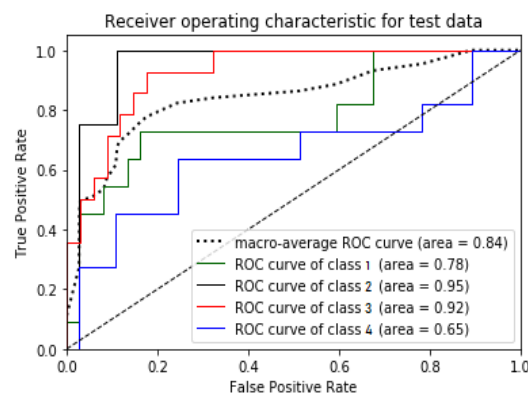


**Figure 11.** Nepal—ROC curves for each class of test dataset. The classifier classified classes 2 and 3 better than classes 1 and 4 samples. A higher false-positive rate appeared for class 4.

Additionally, the area under the curve (AUC) is also used for model performance evaluation as a summarized intelligence model. AUC is the probability of a model to put a positive sample in a higher rank than a negative chosen sample. Therefore, the more substantial area under the curve, the better the model. Table 10 presents the AUC score achieved by each model.
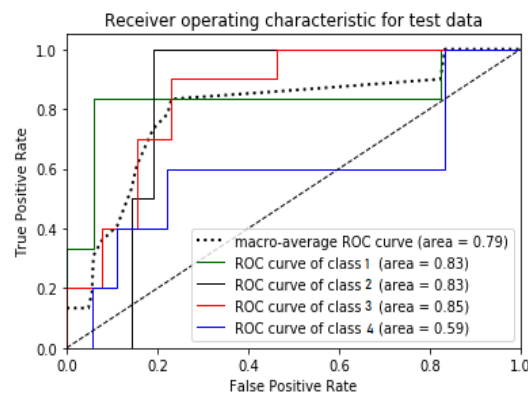
**Figure 12.** South Korea—ROC curves for each class of test dataset. Samples for classes 1 and 3 are identified in a similar trend, whereas class 3 has a maximum true positive rate achieved by the classifier. Several sample data of class 4 are taken as false positive, hence could not fulfill the threshold.

**Table 10.** AUC score of classifiers.

| Classifier | AUC Score (Macro) |
|---|---|
| Ecuador | 0.740 |
| Haiti | 0.840 |
| Nepal | 0.826 |
| South Korea | 0.775 |

## 6. Conclusions

Modern building designs respect the required safety norms while constructing multi-story and intricate architecture, and therefore they possess less risk when subjected to seismic vulnerability. However, existing residential and commercial buildings show several signs of low construction quality, lousy maintenance, and damage signs, further increasing if associated with an earthquake.

This study focused on the modern technique of RVS methods to analyze and minimize the risk factors associated with old and existing buildings. Machine learning is quick and cost-effective when the excellent quality of data is available. RC buildings sample data from four different countries Ecuador, Haiti, Nepal, and South Korea, are interpreted and evaluated using Machine Learning to scrutinize the damage scale for unseen samples. Eight input features, the number of stories, total floor area, column area, concrete wall area (X and Y), masonry wall area (X and Y), and captive columns are considered. The performance of the classifier is evaluated with imbalanced and balanced input data. Balanced data equally distribute the number of feature points among all the classes, by duplicating the data points from minor class, without making any significant change in the information. The classifier showed improved accuracy for each dataset with balanced data. Ecuador, Haiti, and Nepal datasets incurred an accuracy of 60%, 68%, and 67%, respectively, which illustrate that the classifier performed well over the unseen samples. The probable reason behind South Korea's limited performance (less than 50%) could be fewer sample data. A classifier does not significantly evaluate new data if less training data are available, which weakens its learning capacity.

The future study can further extend to evaluate the SVM classifier's performance with more data while examining each feature's influence. The soft margin technique can also be implemented, replacing hyper-parameter tuning to study similar datasets for multi-class classification. Possibly k-fold cross-validation technique can also be used instead of GridSearch.

**Author Contributions:** Conceptualization, E.H. and T.L.; methodology, V.K. and K.J.; validation, T.L. and E.H.; formal analysis, E.H. and V.K.; investigation, K.J. and V.K.; resources, E.H.; data curation, V.K. and K.J.; writing—original draft preparation, E.H., K.J., V.K., and R.R.D.; writing—review and editing, E.H., K.J., V.K., and S.R.; visualization, V.K. and S.R.; supervision, E.H., T.L.; project administration, E.H. and T.L. All authors have read and agreed to the published version of the manuscript.

## Abbreviations

The following abbreviations are used in this manuscript:

| | |
|---|---|
| ACI | American Concrete Institute |
| ANN | Artificial Neural Network |
| AUC | Area under the Curve |
| C | Cost of constraints violation |
| FEMA | Federal Emergency Management Agency |
| ML | Machine Learning |
| N | Number of Stories |
| NRS | Normalized Redundancy Score |
| PI | Priority Index |
| RBF | Radial Basis Function |
| RC | Reinforced Concrete |
| ROC | Receiver Operating Characteristics |
| RVS | Rapid Visual Screening |
| SVM | Support Vector Machine |

## References

1. Gavarini, C. Seismic risk in historical centers. *Soil Dyn. Earthq. Eng.* **2001**, *21*, 459–466. [CrossRef]
2. Jain, S.; Mitra, K.; Kumar, M.; Shah, M. A proposed rapid visual screening procedure for seismic evaluation of RC-frame buildings in India. *Earthq. Spectra-Earthq Spectra* **2010**, *26*, 709–729. [CrossRef]
3. Chanu, N.; Nanda, R. A Proposed Rapid Visual Screening Procedure for Developing Countries. *Int. J. Geotech. Earthq. Eng.* **2018**, *9*, 38–45. [CrossRef]
4. Sinha, R.; Goyal, A. *A National Policy for Seismic Vulnerability Assessment of Buildings and Procedure for Rapid Visual Screening of Buildings for Potential Seismic Vulnerability*; Report to Disaster Management Division; Ministry of Home Affairs, Government of India: Mumbai, India, 2004.
5. FEMA P-154. *Third Edition, Rapid Visual Screening of Buildings for Potential Seismic Hazards: A Handbook*; Homeland Security Department, Federal Emergency Management Agency: Washington, DC, USA, 2015.
6. Harirchian, E. Constructability Comparison Between IBS and Conventional Construction. Ph.D. Thesis, Universiti Teknologi Malaysia, Skudai, Malaysia, 2015.
7. Rai, D.C. *Seismic Evaluation and Strengthening of Existing Buildings*; IIT Kanpur and Gujarat State Disaster Mitigation Authority: Gandhinagar, India, 2005; pp. 1–120.
8. Vallejo, C.B. Rapid Visual Screening of Buildings in the City of Manila, Philippines. In Proceedings of the 5th Civil Engineering Conference in the Asian Region and Australasian Structural Engineering Conference 2010, Sydney, Australia, 8–12 August 2010; p. 513.
9. Mishra, S. *Guide book for Integrated Rapid Visual Screening of Buildings for Seismic Hazard*; TARU Leading Edge Private Ltd.: New Delhi, India, 2014.
10. Luca, F.; Verderame, G. *Seismic Vulnerability Assessment: Reinforced Concrete Structures*; Springer: Berlin/Heidelberg, Germany, 2014. [CrossRef]
11. Chanu, N.; Nanda, R. Rapid Visual Screening Procedure of Existing Building Based on Statistical Analysis. *Int. J. Disaster Risk Reduct.* **2018**, *28*, 720–730. [CrossRef]
12. Özhendekci, N.; Özhendekci, D. Rapid Seismic Vulnerability Assessment of Low- to Mid-Rise Reinforced Concrete Buildings Using Bingöl's Regional Data. *Earthq. Spectra* **2012**, *28*, 1165–1187. [CrossRef]
13. Harirchian, E.; Lahmer, T.; Rasulzade, S. Earthquake Hazard Safety Assessment of Existing Buildings Using Optimized Multi-Layer Perceptron Neural Network. *Energies* **2020**, *13*, 2060. [CrossRef]

14. Arslan, M.; Ceylan, M.; Koyuncu, T. An ANN approaches on estimating earthquake performances of existing RC buildings. *Neural Netw. World* **2012**, *22*, 443–458. [CrossRef]

15. Tesfamariam, S.; Liu, Z. Earthquake induced damage classification for reinforced concrete buildings. *Struct. Saf.* **2010**, *32*, 154–164. [CrossRef]

16. Harirchian, E.; Lahmer, T. Improved Rapid Assessment of Earthquake Hazard Safety of Structures via Artificial Neural Networks. In Proceedings of the 2020 5th International Conference on Civil Engineering and Materials Science (ICCEMS 2020), Singapore, 15–18 May 2020; IOP Conference Series: Materials Science and Engineering; IOP Publishing: Bristol, UK, 2020; Volume 897, p. 012014. [CrossRef]

17. Harirchian, E.; Lahmer, T.; Kumari, V.; Jadhav, K. Application of Support Vector Machine Modeling for the Rapid Seismic Hazard Safety Evaluation of Existing Buildings. *Energies* **2020**, *13*, 3340. [CrossRef]

18. Harirchian, E.; Harirchian, A. *Earthquake Hazard Safety Assessment of Buildings via Smartphone App: An Introduction to the Prototype Features- 30. Forum Bauinformatik: Von jungen Forschenden für junge Forschende: September 2018, Informatik im Bauwesen*; Professur Informatik im Bauwesen, Bauhaus-Universität Weimar: Weimar, Germany, 2018; pp. 289–297. [CrossRef]

19. Harirchian, E.; Jadhav, K.; Mohammad, K.; Aghakouchaki Hosseini, S.E.; Lahmer, T. A Comparative Study of MCDM Methods Integrated with Rapid Visual Seismic Vulnerability Assessment of Existing RC Structures. *Appl. Sci.* **2020**, *10*, 6411. [CrossRef]

20. Ketsap, A.; Hansapinyo, C.; Kronprasert, N.; Limkatanyu, S. Uncertainty and fuzzy decisions in earthquake risk evaluation of buildings. *Eng. J.* **2019**, *23*, 89–105. [CrossRef]

21. Kapetana, P.; Dritsos, S. Seismic assessment of buildings by rapid visual screening procedures. *Wit Trans. Built Environ.* **2007**, *93*, 409–418. [CrossRef]

22. Tesfamariam, S.; Saatcioglu, M. Seismic vulnerability assessment of reinforced concrete buildings using hierarchical fuzzy rule base modeling. *Earthq. Spectra* **2010**, *26*, 235–256. [CrossRef]

23. Şen, Z. Rapid visual earthquake hazard evaluation of existing buildings by fuzzy logic modeling. *Expert Syst. Appl.* **2010**, *37*, 5653–5660. [CrossRef]

24. Harirchian, E.; Lahmer, T. Developing a hierarchical type-2 fuzzy logic model to improve rapid evaluation of earthquake hazard safety of existing buildings. *Structures* **2020**, *28*, 1384–1399. [CrossRef]

25. Harirchian, E.; Lahmer, T. Improved Rapid Visual Earthquake Hazard Safety Evaluation of Existing Buildings Using a Type-2 Fuzzy Logic Model. *Appl. Sci.* **2020**, *10*, 2375. [CrossRef]

26. Cerovecki, A.; Gharahjeh, S.; Harirchian, E.; Ilin, D.; Okhotnikova, K.; Kersten, J. Evaluation of Change Detection Techniques using Very High Resolution Optical Satellite Imagery. *Preface* **2015**, *2*, 20.

27. Valentijn, T.; Margutti, J.; van den Homberg, M.; Laaksonen, J. Multi-Hazard and Spatial Transferability of a CNN for Automated Building Damage Assessment. *Remote. Sens.* **2020**, *12*, 2839. [CrossRef]

28. Morfidis, K.; Kostinakis, K. Seismic parameters' combinations for the optimum prediction of the damage state of R/C buildings using neural networks. *Adv. Eng. Softw.* **2017**, *106*, 1–16. [CrossRef]

29. Sucuoglu, H.; Yazgan, U.; Yakut, A. A Screening Procedure for Seismic Risk Assessment in Urban Building Stocks. *Earthq. Spectra-Earthq Spectra* **2007**, *23*. [CrossRef]

30. Jain, S.; Mitra, K.; Kumar, M.; Shah, M. A rapid visual seismic assessment procedure for RC frame buildings in India. In Proceedings of the 9th US National and 10th Canadian Conference on Earthquake Engineering, Toronto, ON, Canada, 29 July 2010. [CrossRef]

31. Coskun, O.; Aldemir, A.; Sahmaran, M. Rapid screening method for the determination of seismic vulnerability assessment of RC building stocks. *Bull. Earthq. Eng.* **2019**, *18*, 1401–1416. [CrossRef]

32. Askan, A.; Yucemen, M. Probabilistic methods for the estimation of potential seismic damage: Application to reinforced concrete buildings in Turkey. *Struct. Saf.* **2010**, *32*, 262–271. [CrossRef]

33. Morfidis, K.E.; Kostinakis, K.G.; Use Of Artificial Neural Networks in the R/C Buildings' Seismic Vulnerabilty Assessment: The Practical Point. In Proceedings of the 7th International Conference on Computational Methods in Structural Dynamics and Earthquake Engineering, Crete, Greece, 24–26 June 2019.

34. Dritsos, S.; Moseley, V. A fuzzy logic rapid visual screening procedure to identify buildings at seismic risk. *Beton- und Stahlbetonbau* **2013**, 136–143. Available online: https://www.researchgate.net/publication/295594396_A_fuzzy_logic_rapid_visual_screening_procedure_to_identify_buildings_at_seismic_risk (accessed on 30 June 2020).

35. Zhang, Z.; Hsu, T.Y.; Wei, H.H.; Chen, J.H. Development of a Data-Mining Technique for Regional-Scale Evaluation of Building Seismic Vulnerability. *Appl. Sci.* **2019**, *9*, 1502. [CrossRef]

36. Christine, C.A.; Chandima, H.; Santiago, P.; Lucas, L.; Chungwook, S.; Aishwarya, P.; Andres, B. A cyberplatform for sharing scientific research data at DataCenterHub. *Comput. Sci. Eng.* **2018**, *20*, 49. [CrossRef]

37. Sim, C.; Villalobos, E.; Smith, J.P.; Rojas, P.; Pujol, S.; Puranam, A.Y.; Laughery, L.A. Performance of Low-rise Reinforced Concrete Buildings in the 2016 Ecuador Earthquake, Purdue University Research Repository, United States. 2017. Available online: https://purr.purdue.edu/publications/2727/1 (accessed on 10 June 2020).

38. Vera-Grunauer, X. Geer-Atc Mw7.8 Ecuador 4/16/16 Earthquake Reconnaissance Part II: Selected Geotechnical Observations. In Proceedings of the 16th World Conference on Earthquake Engineering (WCEE), Santiago, Chile 9–13 January 2017.

39. O'Brien, P.; Eberhard, M.; Haraldsson, O.; Irfanoglu, A.; Lattanzi, D.; Lauer, S.; Pujol, S. Measures of the Seismic Vulnerability of Reinforced Concrete Buildings in Haiti. *Earthq. Spectra* **2011**, *27*, S373–S386. [CrossRef]

40. *NEES: The Haiti Earthquake Database*; DEEDS, Purdue University Research Repository: Lafayette, IN, USA, 2017. Available online: https://datacenterhub.org/resources/263 (accessed on 2 June 2020).

41. De León, R.O. Flexible soils amplified the damage in the 2010 Haiti earthquake. *Earthq. Resist. Eng. Struct. IX* **2013**, *132*, 433–444. [CrossRef]

42. U.S. *Geological Survey (USGS) 2010*; U.S. Geological Survey: Reston, VA, USA, 2010.

43. Shah, P.; Pujol, S.; Puranam, A.; Laughery, L. *Database on Performance of Low-Rise Reinforced Concrete Buildings in the 2015 Nepal Earthquake*; DEEDS, Purdue University Research Repository: Lafayette, IN, USA, 2015. Available online: https://datacenterhub.org/resources/238 (accessed on 10 June 2020).

44. Tallett-Williams, S.; Gosh, B.; Wilkinson, S.; Fenton, C.; Burton, P.; Whitworth, M.; Datla, S.; Franco, G.; Trieu, A.; Dejong, M.; et al. Site amplification in the Kathmandu Valley during the 2015 M7. 6 Gorkha, Nepal earthquake. *Bull. Earthq. Eng.* **2016**, *14*, 3301–3315. [CrossRef]

45. Sim, C.; Laughery, L.; Chiou, T.C.; Weng, P.W. *2017 Pohang Earthquake—Reinforced Concrete Building Damage Survey*; DEEDS, Purdue University Research Repository: Lafayette, IN, USA, 2018. Available online: https://datacenterhub.org/resources/14728 (accessed on 2 June 2020).

46. Kim, H.S.; Sun, C.G.; Cho, H.I. Geospatial assessment of the post-earthquake hazard of the 2017 Pohang earthquake considering seismic site effects. *ISPRS Int. J. Geo-Inf.* **2018**, *7*, 375. [CrossRef]

47. Smith, E.M.; Mooney, W. A seismic intensity survey of the April 16, 2016 Mw 7.8 Muisne, Ecuador earthquake, and a comparison with strong motion data. *Agu Fall Meet. Abstr.* **2017**, *2017*, S13C–0685.

48. *U.S. Geological Survey (USGS) 2015*; U.S. Geological Survey: Reston, VA, USA, 2015.

49. *U.S. Geological Survey (USGS) 2017*; U.S. Geological Survey: Reston, VA, USA, 2017.

50. Stone, H. Exposure and Vulnerability for Seismic Risk Evaluations. Ph.D. Thesis, University College London, London, UK, 2018.

51. Harirchian, E.; Lahmer, T. Earthquake Hazard Safety Assessment of Buildings via Smartphone App: A Comparative Study. *IOP Conf. Ser. Mater. Sci. Eng.* **2019**, *652*, 012069. [CrossRef]

52. Yakut, A.; Aydogan, V.; Ozcebe, G.; Yucemen, M. Preliminary Seismic Vulnerability Assessment of Existing Reinforced Concrete Buildings in Turkey. In *Seismic Assessment and Rehabilitation of Existing Buildings*; Springer: Dordrecht, The Netherlands, 2003; pp. 43–58.

53. Hassan, A.F.; Sozen, M.A. Seismic vulnerability assessment of low-rise buildings in regions with infrequent earthquakes. *ACI Struct. J.* **1997**, *94*, 31–39.

54. Herrero-Lopez, S. Multiclass support vector machine. In *GPU Computing Gems Emerald Edition*; Elsevier: Burlington, MA, USA, 2011; pp. 293–311.

55. Han, J.; Kamber, M.; Pei, J. 6-mining frequent patterns, associations, and correlations: Basic concepts and methods. In *Data Mining*, 3rd ed.; The Morgan Kaufmann Series in Data Management Systems; Morgan Kaufmann Publishers Inc.: San Francisco, CA, USA, 2012; pp. 243–278.

56. Provost, F. Machine learning from imbalanced data sets 101. In *Proceedings of the AAAI'2000 Workshop on Imbalanced Data Sets*; AAAI Press: Menlo Park, CA, USA, 2000; Volume 68, pp. 1–3.

57. Yücemen, M.; Özcebe, G.; Pay, A. Prediction of potential damage due to severe earthquakes. *Struct. Saf.* **2004**, *26*, 349–366. [CrossRef]

58. Cortes, C.; Vapnik, V. Support-vector networks. *Mach. Learn.* **1995**, *20*, 273–297. [CrossRef]

59. Wang, J.; Yao, Y.; Liu, Z. A new multi-class classification based on non-linear svm and decision tree. In Proceedings of the 2007 Second International Conference on Bio-Inspired Computing: Theories and Applications, Harbin, China, 14–17 September 2007; pp. 117–119.

60. Weston, J.; Mukherjee, S.; Chapelle, O.; Pontil, M.; Poggio, T.; Vapnik, V. Feature selection for SVMs. In *Advances in Neural Information Processing Systems; MIT Press: Cambridge, MA, USA*, 2001; pp. 668–674.

61. Villalobos, E.; Sim, C.; Smith-Pardo, J.P.; Rojas, P.; Pujol, S.; Kreger, M.E. The 16 April 2016 Ecuador Earthquake Damage Assessment Survey. *Earthq. Spectra* **2018**, *34*, 1201–1217. [CrossRef]

62. Goda, K.; Kiyota, T.; Pokhrel, R.M.; Chiaro, G.; Katagiri, T.; Sharma, K.; Wilkinson, S. The 2015 Gorkha Nepal earthquake: Insights from earthquake damage survey. *Front. Built Environ.* **2015**, *1*, 8. [CrossRef]

63. Grigoli, F.; Cesca, S.; Rinaldi, A.P.; Manconi, A.; Lopez-Comino, J.A.; Clinton, J.; Westaway, R.; Cauzzi, C.; Dahm, T.; Wiemer, S. The November 2017 Mw 5.5 Pohang earthquake: A possible case of induced seismicity in South Korea. *Science* **2018**, *360*, 1003–1006. [CrossRef]