POST-EARTHQUAKE BUILDING PERFORMANCE: MACHINE LEARNING FOR 2016–2017 CENTRAL ITALY EARTHQUAKES

A. Aloisio[1], M. M. Rosso[2], L. Coco[1], L. Di Battista[3], B. Di Giacomantonio [4], M. Fragiacomo[1], G. C. Marano[2] & G. Quaranta[5]

[1] Department of Civil, Construction-Architectural and Environmental Engineering, University of L'Aquila, L'Aquila, Italy, angelo.aloisio1@univaq.it

[2] Department of Structural, Geotechnical and Building Engineering, Politecnico di Torino, Turin, Italy

[3] Ministry of Infrastructure and Transport, Public Works Department of Lazio, Abruzzo, and Sardinia, L'Aquila, Italy

[4] Special Office for Reconstruction (Ufficio Speciale Ricostruzione Sisma 2016) - Abruzzo, Teramo, Italy

[5] Department of Structural and Geotechnical Engineering, Sapienza University, Rome, Italy

Abstract: *In post-earthquake situations, the rapid evaluation of the buildings' performances and usability may play a crucial role, especially for an effective emergency management. This research discusses a machine-learning-based workflow to cope with the observed post-earthquake buildings' usability data concerning the 2016–2017 central Italy earthquake sequence. Specifically, the data-driven model is expected to provide an useful supporting tool for predicting the building's usability immediately after an earthquake event. In order to improve the predictive capabilities of the model, the intensity measures have been encompassed in the post-earthquake usability data coming from the rapid evaluation AeDES forms related to the Abruzzi region (Italy). To improve the performances of the intelligent classifiers, a binary classification problem has been formalized to distinguish between no-damaged or slightly damaged structures, i.e. which can be almost immediately occupied again, from those that are partially or entirely unusable, which, on the other hand, require significant restoration interventions or complete demolition. The adopted methodologies for an effective data treatment due to the imbalanced nature of the dataset have been discussed in order to attempt to provide fair training of the models accounting for the minority class. The current study highlighted the main still existing challenges in dealing with a strongly imbalanced dataset likewise the one under investigation. Nevertheless, the promising results emphasized that a properly calibrated machine learning model may provide a useful tool to support the decision-making process in emergency conditions estimating the building performances and usability in post-event seismic scenarios.*

1. Introduction

After seismic events, detailed inspections are crucial for determining the usability and extent of damage to structures, which in turn guides financial estimates for their repair or rebuilding. In Italy, the assessment of building safety and damage uses the AeDES method (Baggio et al., 2007), which categorizes buildings into risk classes from A to F (Del Gaudio, Di Domenico, et al., 2018). Despite such challenges, the accumulated earthquake data over five decades, accessible via the Da.D.O. system, is a critical resource for understanding exposure and vulnerability, contributing to efforts in Disaster Risk Reduction (DRR) by supplying loss data and insights into vulnerability for future seismic activity prediction (Del Gaudio et al., 2017). Additionally, this data is crucial for developing the taxonomies used in evaluation forms (as discussed by Braga et al., 1982; Whitman et al., 1973) and for discerning patterns of vulnerability and damage within earthquake-sensitive areas (Del Gaudio, Ricci, et al., 2018; Drago et al., 2015), which is a key focus of research utilizing these databases to link building features with incurred damages. Despite the comprehensive work on assessing the fragility of Italian structures through post-seismic damage data, there is a research gap concerning the Machine Learning models for predicting the classification outcome, see Nicodemo et al. (2020**)**.

This research uses existing observations from past earthquake-induced damages to predict the seismic performance encoded by the AeEDES forms using machine learning algorithms. This investigation is centered around the sequence of earthquakes that struck Central Italy from 2016 to 2017, reported in Fig.1. The case study uses data on 12,662 buildings gathered via post-earthquake assessments, along with the subsequent structural outcomes. The database compiles a comprehensive set of input variables, inclusive of prevalent measures of seismic intensity. The primary objective is to categorize the buildings into two distinct groups: those that are immediately occupiable or require minimal emergency repairs due to negligible or minor damage, and those that are partially or completely uninhabitable due to extensive earthquake damage. A variety of machine learning strategies are adopted for this categorization process, including K-nearest neighbors, Linear support vector machine, Radial basis function support vector machine, Decision tree, Random forest, Neural network, Adaptive boosting, Naive Bayes, and Quadratic discriminant analysis. To address the challenge of an unbalanced input database, two preprocessing methods are applied: Principal component analysis and Synthetic minority oversampling technique. Metrics such as precision, recall, F1-score, and accuracy are used to assess the efficacy of the machine learning models developed. The findings prove that an optimally calibrated machine learning model can efficiently deliver credible predictions concerning the performance and occupancy suitability of buildings post-earthquake, relying on minimal data that can be readily acquired.

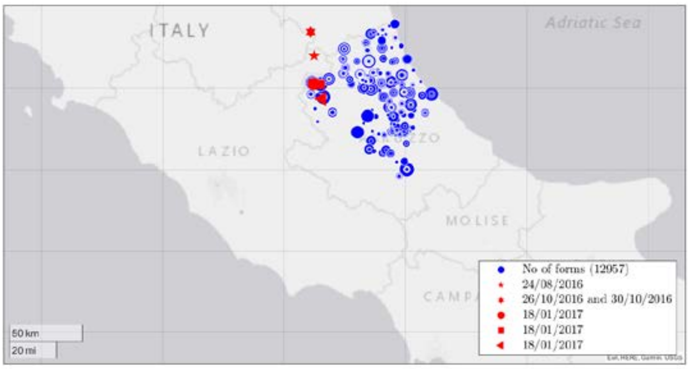


Figure 1. Epicentre of the most significant seismic events of the 2016-2017 Central Italy seismic sequence.

1.1. The AeDES form

The AeDES form consists of nine sections spread across three pages, with an extra fourth page dedicated to explanatory notes.

Section 1 - BUILDING IDENTIFICATION provides survey details and building information.

Section 2 - BUILDING DESCRIPTION offers a concise overview of the structure, covering factors like total storeys (including basements), storey height, surface area, construction age, and usage.

Section 3 - TYPOLOGY focuses on structure characteristics, particularly emphasizing masonry buildings.

Section 4 - DAMAGE TO STRUCTURAL ELEMENTS catalogs damage observed, structured by location and intensity.

Section 5 - DAMAGE TO NON-STRUCTURAL ELEMENTS records non-structural damage, similarly organized by location and intensity.

Section 6 - OUTSIDE DANGER signals potential hazards from neighboring structures, networks, slopes, and executed emergency measures.

Section 7 - SOIL AND FOUNDATIONS outlines site morphology and foundation instability likelihood.

Section 8 - RISK CLASS denotes the assigned risk category.

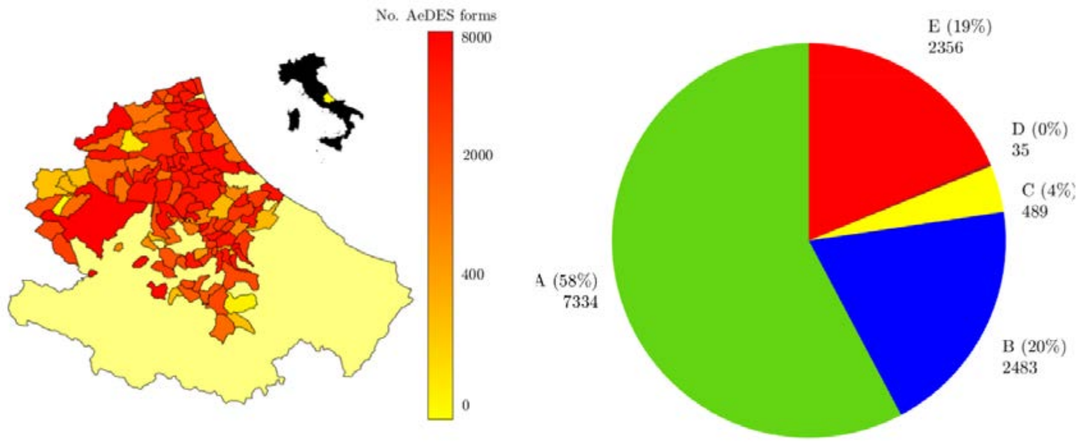
Section 9 - OTHER REMARKS comprises a blank page for evaluator notes.

It's crucial to highlight that while classes A, B, C, and E directly correlate with damage severity, signifying increasing severity from A to E, classes D and F don't directly indicate damage levels. Class D implies the need for a subsequent survey, making its prediction solely based on the damage matrix irrelevant as it depends on other factors. On the other hand, class F represents external risk, meaning it's not directly linked to the surveyed building's observed damage but rather to nearby building damage, which might impact the surveyed building's safety. Consequently, the damage-related classes are A, B, C, and E.

Furthermore, it's important to recognize that the risk classes, resulting from the AeDES form, indeed relate to the maximum state reimbursement for building reconstruction or repair, as mentioned earlier. However, it's crucial to acknowledge that this direct correlation between AeDES outcomes and funding was specific to the L'Aquila earthquake. Funding processes varied for subsequent earthquakes, such as Emilia, Centre Italy, and Ischia. In these later events, an additional parameter called the operational level factored into the conventional parametric cost assessment within each risk class provided by AeDES forms.

1.2. Analysis of the dataset

Data from 12,662 buildings located in Abruzzo collected after the third seismic sequence are processed in this study, see Fig.2.



(a) (b)

Figure 2. (a) Number of AeDES forms available for each municipality within the considered study; (b) Consequence class assignment to the buildings surveyed through AeDES forms within the study.

The building stock primarily consists of low-rise structures, with 30% having two floors and 40% having three floors. A significant proportion of the buildings (76%) were constructed before 1975, suggesting older construction practices and possibly outdated seismic design. There has been a decline in new building constructions since the 1980s. Most buildings (91%) have a floor area between 50 m² to 200 m², are isolated (38%), and have an average inter-storey height between 2.50 m and 3.49 m (86%). The majority of the buildings (80%) are masonry, and about 20% have a concrete structure. The roofs are mainly non-pushing, either heavy or light (70% collectively), with a substantial number still having pushing roofs. In terms of masonry quality, buildings generally have a regular texture, but a considerable number have irregular textures. There are more than twice as many masonry buildings with chains or curbs compared to those without, and isolated pillars are uncommon (8%). There are only a few hybrid or reinforced masonry buildings.Regarding topography, plane site is uncommon, with over 80% of the buildings located on sloped terrain, indicating mountainous locations.Analyzing the proximity to the epicenter, over 80% of the buildings are within 40 km of the earthquake's epicenter, which suggests a high exposure to seismic risk.The consequence class distribution, as determined by AeDES forms, reveals an imbalance. About 60% of buildings are in class A, meaning they were accessible post-earthquake. On the other hand, approximately 20% fall into classes B (temporarily unusable) or E (permanently unusable). Only 4% are in the intermediate class C, and a very few in class D (not conclusive outcomes). The data reveal a region with a predominance of older, low-rise masonry buildings, many of which have shown vulnerability to seismic activity. There is a clear need for targeted seismic retrofitting and reinforcement, especially given the proximity of much of the building stock to earthquake epicenters. Furthermore, the imbalances and missing data points in the dataset underscore the necessity for a more systematic data collection and assessment approach to better understand the seismic resilience of the built environment in Abruzzo.

2. ML models for binary classification of building usability

2.1. Database pre-processing

A binary classification problem has been formalized to distinguish between no-damaged (A) or slightly damaged structures (B), i.e. which can be almost immediately occupied again, from those that are partially (C) or entirely unusable (E), which, on the other hand, require significant restoration interventions or complete demolition. With a deeper inspection of the dataset, it is important to notice that some data are missing, thus those rows were excluded reducing the database to 12.063 samples. The dataset was split into a training and test set with the hold-out method with proportions of 85% and 15% respectively, considering a stratified sampling scheme to observe the relative imbalance nature of the data. Since the post-earthquake building performance strongly depends on the nature of the action that struck the building, some well-acknowledged intensity measures of the most important seismic event in the earthquake sequence have been encompassed within the dataset. Specifically, the added features are the period of vibration, horizontal and vertical peak ground acceleration (PGA_H and PGA_V), the horizontal pseudo-spectral acceleration s_{a_H}, the Arias and Housner intensities, and the duration of the seismic event. Therefore, the input variables of the ML models are reported in Tab.1. In detail, the explanatory features of the binary classification problem are composed of 13 categorical variables collected from the AeDES forms, and 9 continuous variables related to the input seismic event characterization. All these features have been thus encoded and scaled into a proper numerical information format to feed the ML algorithms.

Table 1: Database of input explanatory variables and outcome.

No.	Input variable	data type	No.	Input variable	data type
1	Building position	Categorical	14	Soil category	Categorical
2	No. of floors	Discrete numerical	15	Epicentral distance	Continous numerical
3	Inter-storey height	Discrete numerical	16	Period of vibration	Continous numerical
4	Avarage floor surface	Discrete numerical	17	PGA_H [m/s^2]	Continous numerical
5	Age	Discrete numerical	18	PGA_V [m/s^2]	Continous numerical
6	Typology	Categorical	19	s_{a_H} [m/s^2]	Continous numerical
7	Roof	Categorical	21	Arias Intensity	Continous numerical
8	Masonry texture	Categorical	22	Housner Intensity	Continous numerical
9	Curbs/chains	Categorical	23	Duration [s]	Continous numerical
10	Isolated pillar	Categorical	24	Classification score	Categorical
11	Hybrid masonry	Categorical			
12	Reinforced masonry	Categorical			
13	Site morphology	Categorical			

2.2. Results and discussion

As previously stated, a binary classification problem has been set in which the two classes are related to the merged building performance scores AB (immediately or almost immediately usable buildings) versus the building performances with scores CE (partially or entirely unusable buildings). A set of various state-of-art ML learners have been employed to cope with this task, thus including the K-nearest neighbors (KNN), the Linear Support Vector Machine (LSVM), the Decision Tree (DT) algorithm, the Random Forest (RF) algorithm, and artificial neural networks with the Multi-Layer Perceptron (MLP) architecture.

To effectively deal with the imbalanced nature of the dataset, the authors employed one of the most widespread and acknowledged techniques within the Machine Learning field for imbalance learning, i.e. the Synthetic minority over-sampling technique (SMOTE), see e.g. Haibo et al. (2013). The SMOTE algorithm attempts to provide a more reliable scheme for the oversampling trying to reduce the risk of overfitting issues related to a naïve oversampling method composed of mere random duplications of existing samples belonging to the minority class. In particular, it artificially attempts to reproduce new and non-replicated samples for the minority class by reproducing the similarities of actual samples. The new instances are therefore synthetic because they are produced by extrapolating the similarity information from existing samples of the minority class, as explained in Chawla et al. (2002).

Another crucial aspect to take into careful consideration when dealing with an imbalanced dataset is the adoption of proper and unbiased evaluation metrics. The classification accuracy is widely recognized nowadays being a poor informative metric not reflecting at all the real learned skills of the ML models, see Haibo et al. (2013). Therefore, to reflect the actual performances of a ML model which effectively learned hidden patterns accounting also for the minority class, precision and recall metrics were identified as the most promising metrics for this task among the others, see Haibo et al. (2013). Moreover, to provide a unique synthetic metric, special attention gained the balance accuracy, which represents the average of the recall on the two classes.

Table 2: ML results on binary classification problem AB vs CE.

ML Algorithm	Evaluation Metrics	Strategy			
		No SMOTE		SMOTE	
		Class AB	Class CE	Class AB	Class CE
KNN	Accuracy	75.0		67.2	
	Balanced accuracy	54.6		60.4	
	Precision	79.3	37.8	82.9	33.8
	Recall	91.7	17.5	72.8	48.0
Linear SVM	Accuracy	62.3		64.8	
	Balanced accuracy	65.9		65.3	
	Precision	88.2	34.0	86.8	35.0
	Recall	59.4	72.4	64.4	66.3
DT	Accuracy	60.9		64.4	
	Balanced accuracy	64.8		64.9	
	Precision	87.7	33.0	86.6	34.6
	Recall	57.8	71.9	64.0	65.8
RF	Accuracy	64.1		67.7	
	Balanced accuracy	65.8		64.7	
	Precision	87.4	34.8	85.6	36.5
	Recall	62.8	68.7	70.1	59.4
MLP	Accuracy	76.4		69.2	
	Balanced accuracy	53.9		63.1	
	Precision	79.0	41.4	84.2	36.8
	Recall	94.7	13.1	74.1	52.0

The training phase of the ML models has been performed using the stratified k-fold cross-validation with a number of folds $k = 10$. This cross-validation scheme permitted the various folds accounting for the actual distribution of samples between the majority and minority classes. Afterward the training phase, the hold-out test set has been adopted to compute the classification performance of the trained models using previously unseen data. The classification metrics have been reported in Tab.2. The results evidenced in general that the classification improvements were modest with the SMOTE technique, but required a higher computational effort. Therefore, in this case, the results in Tab.2 demonstrated that better results can be obtained without adopting specific techniques to restore the imbalanced nature of the dataset prior to the learning phase.

Furthermore, it is worth highlighting that the most accurate models are often obtained by means of machine learning algorithms that exhibit the poorest performance in terms of accuracy, thus reflecting the poor effectiveness of using this metric. In fact, focusing on the minority class specifically, the LSVM with standard strategy without applying the SMOTE oversampling scheme exhibited the highest absolute value of balanced accuracy (65.9%), as well as, recall on the minority class. The LSVM is immediately followed by the RF in terms of balanced accuracy (65.8% without the SMOTE technique). It is worth noting that the LSVM better captures the minority class with a recall of 72.4% on class CE versus a recall of 59.4% on class AB. On the other hand, the RF learned to capture in an almost equilibrated manner both classes, with a recall for class CE and AB respectively of 68.7% and 62.8%.

3. Conclusions

In this research, the authors proposed a potential machine-learning-based workflow to process the observed post-earthquake buildings' usability data concerning the 2016–2017 central Italy earthquake sequence. In particular, the authors adopted the building performance database coming from the rapid evaluation of AeDES forms related to the Abruzzi region (Italy). A binary classification problem has been formulated to predict and discriminate the buildings' performance between usable and unusable buildings. The earthquake data have been encompassed in the AeDES database considering some acknowledged earthquake intensity measures. A set of various state-of-art ML learners have been employed to cope with this task, thus including the K-nearest neighbors (KNN), the Linear Support Vector Machine (LSVM), the Decision Tree (DT) algorithm, the Random Forest (RF) algorithm, and artificial neural networks with the Multi-Layer Perceptron (MLP) architecture. Moreover, to deal with the imbalanced nature of the dataset, the results of a pre-processing performed with the Synthetic minority over-sampling technique (SMOTE) have been compared with the resulting metrics without applying the SMOTE method. In this specific case, the results demonstrated that the positive effects of the SMOTE were quite limited. Therefore the best classification performances were obtained without SMOTE, and only with a stratified k-fold cross-validation scheme. The LSVM and the RF algorithms presented the best-balanced accuracy metrics, highlighting a better capability to capture the minority class for the LSVM against the majority one. On the other hand, the RF showed more equilibrated performances. In conclusion, machine learning tools may provide in the next future some reliable and essential tools to support the delicate decision-making process in emergency conditions such as within post-event seismic scenarios. However, there are still topical challenges to get through and much research effort may be directed toward the imbalance learning, which is a typical situation in the Civil Engineering field.

4. Acknowledgments

This study was carried out within the PRIN 2022 project entitled ``Artificial Intelligence for SUstainable seismic risk reduction of STructures (AI-SUST)" (project code: 2022LEFKHS) –funded by the Ministero dell'Università e della Ricerca –within the PRIN 2022 program (D.D.104 -02/02/2022). This manuscript reflects only the authors' views and opinions and the Ministry cannot be considered responsible for them. Marco Martino Rosso, Giuseppe Carlo Marano, and Giuseppe Quaranta acknowledge the support received through the PRIN project ``Artificial Intelligence for SUstainable seismic risk reduction of STructures (AI-SUST)" (project code: 2022LEFKHS) funded by the Italian Ministry of University and Research.

5. References

Aloisio, A., Rosso, M.M., De Leo, A., Fragiacomo, M., & Basi, M. (2023). Damage classification after the 2009 L'Aquila earthquake using multinomial logistic regression and neural networks. International Journal of Disaster Risk Reduction.

Baggio, C., Bernardini, A., Colozza, R., Corazza, L., Della Bella, M., Di Pasquale, G., Dolce, M., Goretti, A., Martinelli, A., Orsini, G., Papa, F., & Zuccaro, G. (2007). Field manual for post-earthquake damage and safety assessment and short term countermeasures (AeDES). *Field Manual for Post-Earthquake Damage and Safety Assessment and Short Term Countermeasures (AeDES)*.

Braga, F., Dolce, M., & Liberatore, D. (1982). A statistical study on damaged buildings and an ensuing review of the MSK-76 scale. *Proceedings of the Seventh European Conference on Earthquake Engineering*, 431–450.

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: synthetic minority over-sampling technique. Journal of artificial intelligence research, 16, 321-357.

Del Gaudio, C., De Martino, G., Di Ludovico, M., Manfredi, G., Prota, A., Ricci, P., & Verderame, G. M. (2017). Empirical fragility curves from damage data on RC buildings after the 2009 L'Aquila earthquake. *Bulletin of Earthquake Engineering*, *15*(4), 1425–1450. https://doi.org/10.1007/s10518-016-0026-1

Del Gaudio, C., Di Domenico, M., Ricci, P., & Verderame, G. M. (2018). Preliminary prediction of damage to residential buildings following the 21st August 2017 Ischia earthquake. *Bulletin of Earthquake Engineering*, *16*, 4607–4637.

Del Gaudio, C., Ricci, P., & Verderame, G. M. (2018). A class-oriented mechanical approach for seismic damage assessment of RC buildings subjected to the 2009 L'Aquila earthquake. *Bulletin of Earthquake Engineering*, *16*, 4581–4605.

Drago, C., Ferlito, R., & M, Z. (2015). Clustering of damage variables for masonry buildings measured after L'Aquila earthquake. *SIS: Statistics and Demography*.

Haibo, H., & Yunqian, M. (2013). Imbalanced learning: foundations, algorithms, and applications. Wiley-IEEE Press, 1(27), 12.

Nicodemo, G., Pittore, M., Masi, A., & Manfredi, V. (2020). Modelling exposure and vulnerability from post-earthquake survey data with risk-oriented taxonomies: AeDES form, GEM taxonomy and EMS-98 typologies. *International Journal of Disaster Risk Reduction*, *50*, 101894.

Whitman, R. V, Reed, J. W., & Hong, S. T. (1973). Earthquake damage probability matrices. *Proceedings of the Fifth World Conference on Earthquake Engineering*, 2531–2540.