Multi-objective Multi-layer Perceptron Architecture Optimization for Rapid Post-earthquake Damage Evaluation

Original
Multi-objective Multi-layer Perceptron Architecture Optimization for Rapid Post-earthquake Damage Evaluation / Rosso, M. M.; Aloisio, A.; De Leo, A. M.; Basi, M.; Cirrincione, G.; Marano, G. C. (SMART INNOVATION, SYSTEMS AND TECHNOLOGIES). - In: Advanced Neural Artificial Intelligence: Theories and Applications[s.l] : Springer, 2025. - ISBN 9789819609932. - pp. 95-105 [10.1007/978-981-96-0994-9_9]

Availability:
This version is available at: 11583/3001232 since: 2025-06-24T08:28:52Z

Publisher:
Springer

Published
DOI:10.1007/978-981-96-0994-9_9

Multi-objective multi-layer perceptron architecture optimization for rapid post-earthquake damage evaluation

Marco M. Rosso[1,*] [0000−0002−9098−4132], Angelo Aloisio[2] [0000−0002−6190−0139], Andrea Matteo De Leo[2], Maria Basi[3], Giansalvo Cirrincione[4] [0000−0002−2894−4164], and Giuseppe C. Marano[1,5] [0000−0001−8472−2956]

[1] Politecnico di Torino, DISEG, Department of Structural, Geotechnical and Building Engineering, Corso Duca Degli Abruzzi, 24, Turin 10128, Italy,
Email: marco.rosso@polito.it, giuseppe.marano@polito.it
[2] Civil Environmental and Architectural Engineering Department, Università degli Studi dell'Aquila, via Giovanni Gronchi n.18, L'Aquila, 67100, Italy
Email: angelo.aloisio1@univaq.it
[3] Abruzzo Region Risk Prevention of Civil Protection, L'Aquila, Italy
[4] University of Picardie Jules Verne, Lab. LTI, Amiens, France,
Email: exin@u-picardie.fr
[5] College of Civil Engineering, Fuzhou University, Fuzhou, Fujian 350108, China
* Corresponding Author

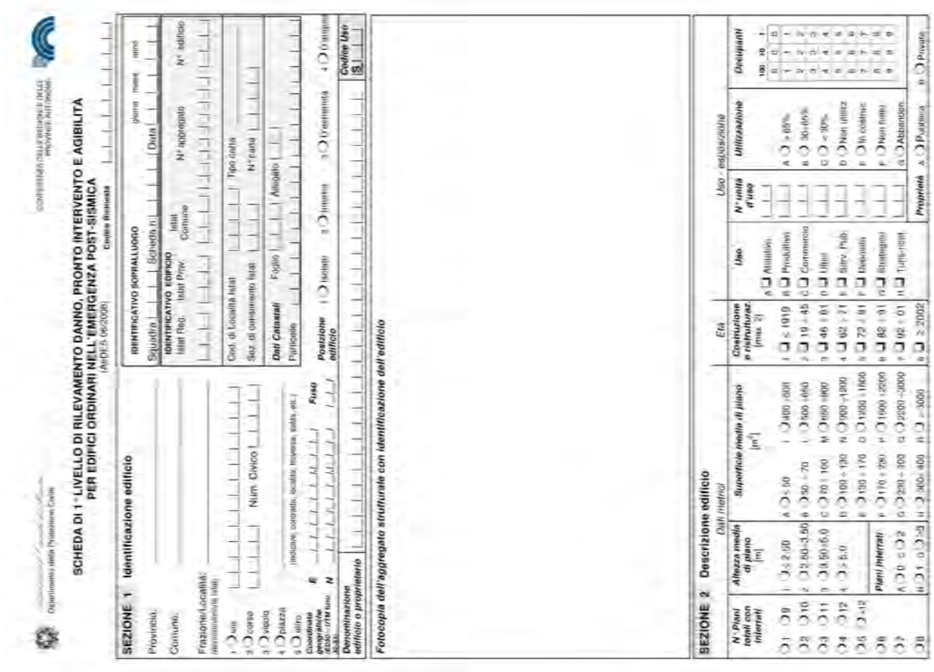
Abstract. Immediately after a seismic crisis, engineers are demanded to provide rapid post-earthquake damage evaluations to globally assess economic losses, residual overall serviceability, and the general safety conditions of the hit existing heritage. Since 1997 and subsequent amendments, in Italy, the inspectors must compile the AeDES forms. Composed of nine sections, the form-filling provides about sixty categorical features to aid the inspector to elaborate a categorical judgment of the rapid qualitative safeness of the structure. This judgment is categorized into six damage classes denoted with capital letters from A to F designating, respectively, fully usable buildings until condemned constructions. Despite the AeDES forms' features should help the inspector to elaborate a quite fair judgment, there is still a certain level of subjectiveness, since there is not any strictly underlining model to objectively convey these sixty categorial features unequivocally toward a specific judgment. Therefore, in the current study, the authors analyzed the AeDES forms data coming from 878 public school buildings, thus exploring the possibility to aid the inspector with a multi-layer perceptron (MLP) neural network model. Precisely, the AeDES forms under investigation are referred to the L'Aquila city seismic event which hit the Abruzzi region of central Italy in 2009. Considering the sensitivity of the current multinomial classification performances in the presence of limited data regarding the MLP architecture topology, the authors formalized a multi-objective optimization problem to find an optimal architecture with the minimum number of hidden neurons and simultaneously maximize the classification accuracy.

Keywords: Multi-objective optimization · Artificial Neural Network · Multi-Layer Perceptron · Post-earthquake Damage Evaluation · Earthquake Engineering · Post-earthquake Assessment

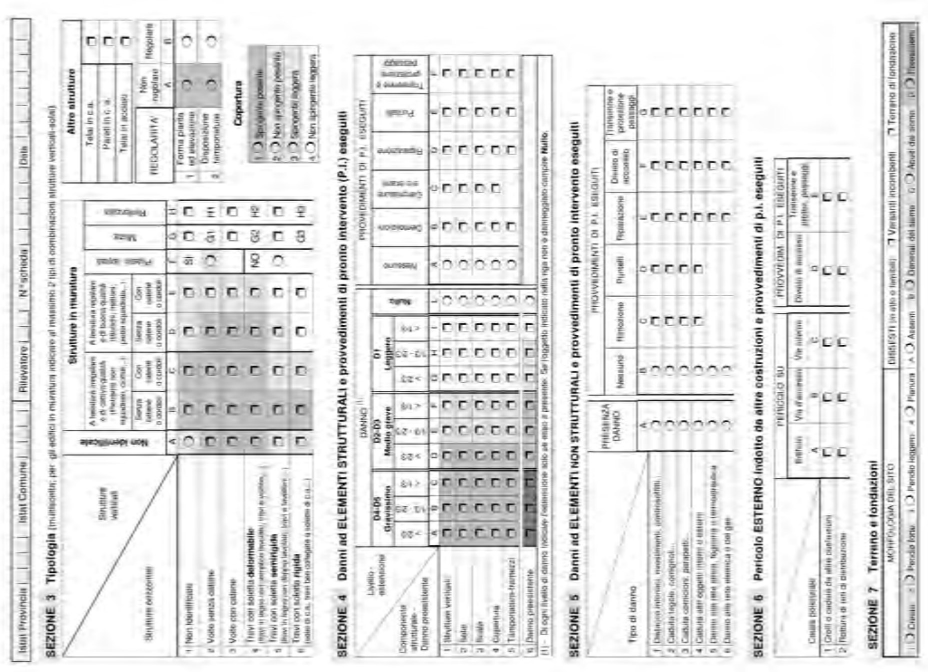
1 Introduction

The Italian Rapid Post-Earthquake Damage evaluation forms, also denoted as AeDES forms (*Agibilitá e Danno nell'Emergenza Sismica*) are composed of 9 sections that are filled out by a structural engineering in charge of conducting the post-earthquake assessment of a building [1,2]. An example of the AeDES forms is provided in Figure 1. The compiled matrix in sections from 1 to 7 delivers 60 categorical features which are functional for the inspector to elaborate a categorical judgment of the safeness in section 8 of the structure under study after the seismic event. Specifically, the possible output judgments are six: class A means that the building is accessible; class B denotes a temporarily unusable building; class C indicates a partially unusable building; class D refers to a temporarily unusable building that should be re-examined; class E is used for a unusable building; class F denotes an unusable building due to external risk only. In the current study, the authors analyzed the AeDES forms data coming from 878 public school buildings, thus exploring the possibility to aid the inspector with a multi-layer perceptron (MLP) neural network model. Precisely, the AeDES forms under investigation are referred to the L'Aquila city seismic event which hit the Abruzzi region of central Italy in 2009. An overview of the current dataset is reported in Figure 2.

Considering the sensitivity of the current multinomial classification performances in the presence of limited data regarding the MLP architecture topology, the authors formalized a multi-objective optimization problem to find an optimal architecture with the minimum number of hidden neurons and simultaneously maximize the classification accuracy. The architecture of the artificial neural network (ANN) should be carefully evaluated, since it strongly affects the model complexity, the classification accuracy rate, and the model's tendency to underfitting or overfitting issues [3,4]. Specifically, to minizime the likelihood of overfitting issues, it is advisable to reduce at most the number of learnable parameters compared to the number of limited data at disposal, however attempting to avoid underfitting issues at the same time. In the literature, various attempts have been already provided to define an optimal architecture for ANN as a single-objective optimization problem [4,5,6,7,8,9]. In [8] the objective function has been set as the squared error between the MLP output and the desired output, and posing a regularization term to control the learnable weights during training. Conversely, the constraints were formulated to guarantee the existence of at least one hidden layer and also to ensure all neurons are fully-connected between two consecutive layers. Evolutionary algorithms (EAs) have been extensively adopted in the existing literature to optimize the ANN architecture, such as the particle swarm optimization [10], the ant

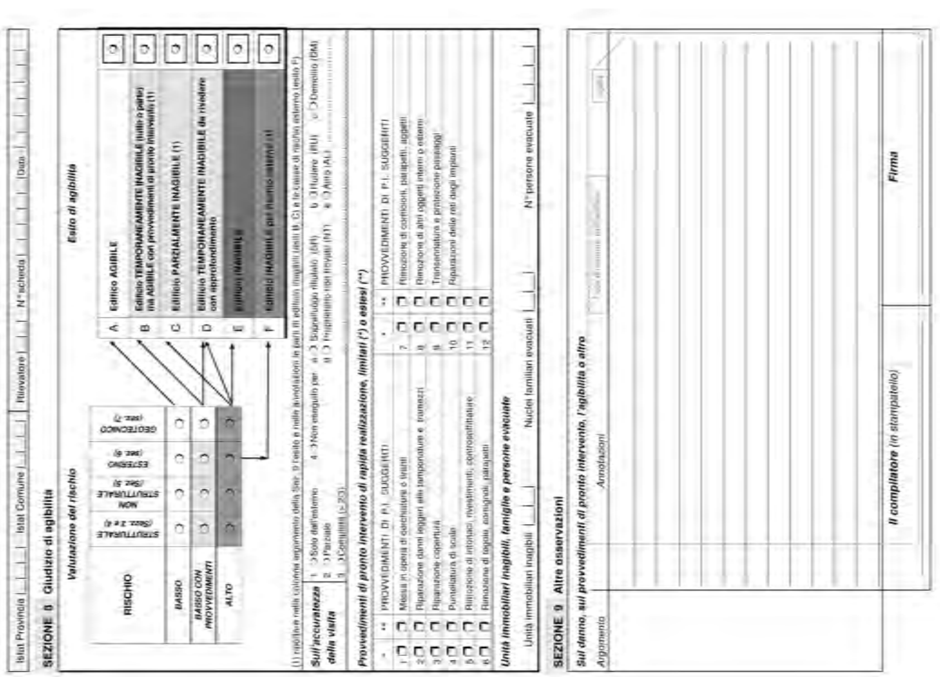


(a)



(b)

Fig. 1. Italian AeDES form (06/2008).



(c)

Fig. 1. Italian AeDES form (06/2008) (cont.).

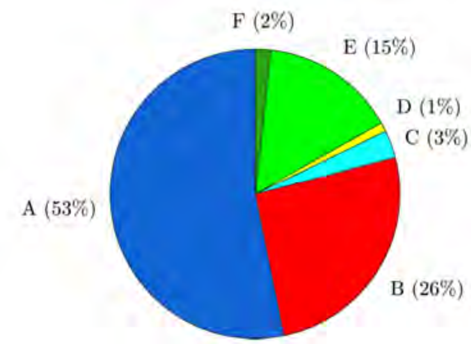


Fig. 2. AeDES forms dataset overview.

colony optimization [11], the simulated annealing in [12] among others. Nonetheless, the most widely employed EA within this context is the genetic algorithm (GA) formulated by Jonhn Holland in the '70s [13]. GA has been adopted e.g. in [8,14]. Alternative approaches to study the optimal architecture of a MLP were early proposed by Hassibi and Stork [15] with optimal brain surgeon (OBS) strategy. Without any apparent restrictive assumption, the OBS strategy leverages information from the second order derivative of the error function to perform a network pruning in order to find the best architecture topology [16]. In [17] MLP architecture information has been encoded with a directed graph modeling. Few efforts of using a MLP architecture multi-objective optimization approach have been also provided e.g. in [18].

The current document is organized as follows. Section 2 illustrates the applied methodology for the AeDES forms classification, whose major findings and results are reported in section 3.

2 Methodology

In the current study, the authors analyzed the AeDES forms data coming from 878 public school buildings, thus exploring the possibility to aid the inspector with a multi-layer perceptron (MLP) neural network model. The authors selected only those output judgments that actually represent a damage class, thus limiting the multi-class classification to four output classes: A, B, C, and E. The authors analyzed the optimal MLP architecture according to a multi-objective optimization strategy. Specifically, the design vector $\boldsymbol{x}$ defined in the admissible search space range Ω, encodes the architecture information. Precisely, the length of the vector defines the number of hidden layers, whereas each vector's components are the number of neuron units existing in every hidden layer. After a preliminary evaluation, the authors set the same hyperbolic tangent activation function to every hidden layer, while

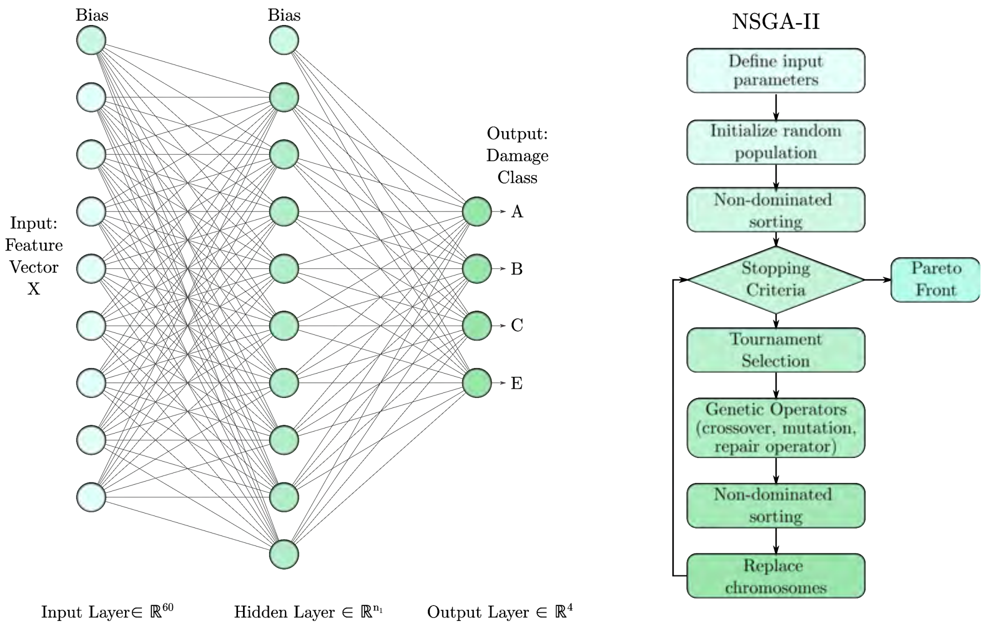


Fig. 3. NSGA-II algorithm workflow used to tackle the MLP optimal architecture multi-objective optimization problem.

posing a softmax function to the output layer accounting for the multi-class classification task under investigation. Furthermore, two contrasting objective functions were defined. On one side, the architecture should be selected in order to provide the best classification accuracy $\alpha_{[\%]}$, whereas, on the other side, it is desirable to diminish the number of hidden layers and the number of units, i.e. the number of learnable weights, in order to reduce the model's complexity. Therefore, accounting for an integer-valued design vector $\boldsymbol{x}$ whose components varies between 0 and 100, and considering the classification error $\varepsilon_{[\%]}$ as the complement to the accuracy, the multi-objective optimization problem statement is defined as below:

$$\min_{\boldsymbol{x}\in\Omega} \boldsymbol{f}(\boldsymbol{x}) = \begin{cases} \min_{\boldsymbol{x}} f_1(\boldsymbol{x}) = \varepsilon_{[\%]} = 100 - \alpha_{[\%]}(\boldsymbol{x}) \\ \min_{\boldsymbol{x}} f_2(\boldsymbol{x}) = \sum_{j=1}^{k} x_j \end{cases} \tag{1}$$

The above-mentioned double minimization problem has been processed with the acknowledged non-dominated sorting genetic algorithm (NSGA-II) [19] using an existing Python implementation [20]. Figure 3 illustrates the main workflow of the NSGA-II algorithm. A 5-fold cross-validation approach has been implemented to guarantee sufficient generalization of every candidate analyzed during the NSGA-II procedure. Thus, the accu-

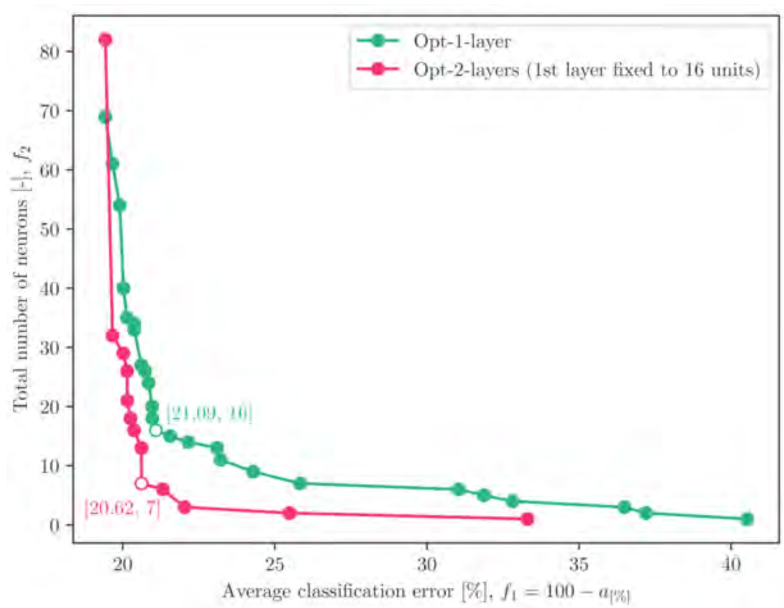


Fig. 4. Pareto front chart of the non-dominated solutions found by the NSGA-II algorithm for the MLP architecture multi-objective optimization problem.

racy term $\alpha_{[\%]}$ in Eq.(1) is computed as the average accuracy among the 5 models trained for the cross-validation method.

3 Results and discussions

At first, the authors analyzed the optimization problem considering an arbitrary number of five possible hidden layers. However, the optimal solution does not remove any hidden layer, but unexpectedly left some layers toward the output one with a single neuron. Using a hidden layer with a single unit has been interpreted as a simple, but possibly useless, transformation of the input hidden states to the subsequent one. Therefore, to better control the complexity growth of the model, the authors manually imposed the number of hidden layers starting from a single one. A Pareto front of non-dominated solutions has been found and it is depicted by the piece-wise green line of Figure 4. The optimum has been assumed in the knee-point of the curve, thus finding that a single hidden layer with 16 units represents the trade-off optimum. In fact, the average classification error $\varepsilon_{[\%]}$ is equal to 21.09%, so that the overall average accuracy is equal to 78.91%.

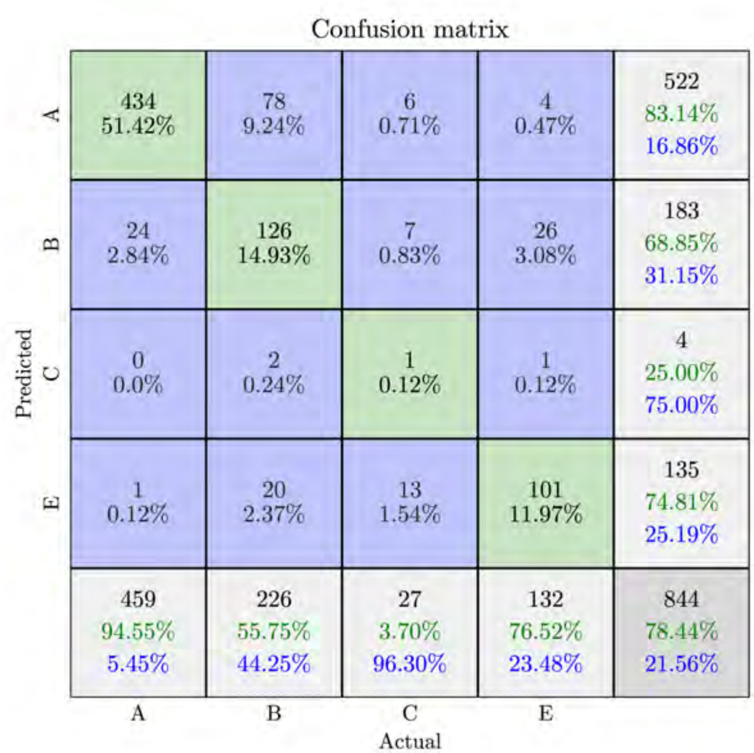


Fig. 5. Average confusion matrix of the optimal ANN architecture with a single hidden layer and 16 hidden neurons.

Thereafter, to progressively increase the model complexity in a controlled way, the authors analyzed an MLP with two hidden layers while imposing the first hidden layer on the previously found optimum number of neurons equal to 16. A new Pareto front of non-dominated solutions has been found and it is depicted by the piece-wise magenta line of Figure 4. In this case, the optimal trade-off number of units for the second hidden layer is equal to 7 at the knee point of the Pareto front. In this case, the average classification error $\varepsilon_{[\%]}$ slightly decreased to 20.62% than before, thus delivering an overall average accuracy of 79.38%.

It is worth noting that passing from one to two hidden layers a significant relative increase of the model complexity has been recorded, although the overall accuracy increased modestly. Therefore, further addition of hidden layers appears quite unreasonable, and therefore the best MLP architecture with the lower complexity can be considered as composed of one single hidden layer with 16 units. Finally, the confusion matrix in Figure 5 illustrates the average classification performances from the $k = 5$ folds cross-validation procedure of the optimized single hidden layer ANN model.

From a visual inspection of the confusion matrix, it is evident that the overall accuracy stalls at about 80%. Furthermore, the MLP model seems to struggle particularly with class C. The main reason may be due to the great unbalance of class C with respect to the others, as demonstrated in Figure 2. Moreover, it is evident from the confusion matrix how class B tends to be partially confused with class A. A feasible way to further improve the classification accuracy may be proposing a binary classification only with class A on one side, and a damage class on the other side, thus encapsulating classes B, C, and E.

4 Conclusions and future developments

In the current study, the authors analyzed the Italian post-earthquake rapid evaluation forms data, i.e. the AeDES forms, coming from 878 public school buildings, thus exploring the possibility to aid the inspector with a multi-layer perceptron (MLP) neural network model. Precisely, the AeDES forms under investigation are referred to the L'Aquila city seismic event which hit the Abruzzi region of central Italy in 2009. Despite six possible outputs are possible from the AeDES forms, the authors selected only those output classes that actually represent damage classes, thus limiting the multi-class classification to four output classes (A, B, C, and E). The optimal MLP architecture problem has been solved through the meta-heuristic non-dominated sorting genetic algorithm (NSGA-II). Finally, the best-found trade-off model with an acceptable global classification accuracy of about 80% and a contained model complexity has been located at the knee point of the Pareto front, thus composed of one single hidden layer with 16 units. In conclusion, the confusion matrix in Figure 5 evidenced how class C was the most unsettled for the classifier model, probably due to its significant intrinsic unbalance compared to the other classes. Future studies may involve the dimensionality reduction of the feature database, e.g. according to clustering techniques or manifold learning methods. This may help new representations and probably more discriminative features to improve the machine learning model's classification ability.

Acknowledgments

The authors acknowledge the Department of Civil Protection of the Abruzzo Region for providing us with the data used in this investigation. Giuseppe Carlo Marano and Marco Martino Rosso acknowledge the support from PNRR MUR project PE0000013-FAIR (National Recovery and Resilience Plan – NRRP, Mission 4, Component 2, Investment 1.3, NextGenerationEU – D.D. 341 15/3/2022 and D.D. 1555 11/10/2022, PE0000013).

References

1. 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.: 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) (2007), cited By 181
2. Marano, G.C., Rosso, M.M., Aloisio, A., Cirrincione, G.: Generative adversarial networks review in earthquake-related engineering fields. Bulletin of Earthquake Engineering pp. 1–52 (2023)
3. Aggarwal, C.C., et al.: Neural networks and deep learning. Springer 10(978), 3 (2018)
4. Aloisio, A., Contento, A., Alaggio, R., Quaranta, G.: Physics-based models, surrogate models and experimental assessment of the vehicle–bridge interaction in braking conditions. Mechanical Systems and Signal Processing 194, 110276 (2023)
5. Domashova, J.V., Emtseva, S.S., Fail, V.S., Gridin, A.S.: Selecting an optimal architecture of neural network using genetic algorithm. Procedia Computer Science 190, 263–273 (2021)
6. Pham, T.A., Tran, V.Q., Vu, H.L.T., Ly, H.B.: Design deep neural network architecture using a genetic algorithm for estimation of pile bearing capacity. PLoS One 15(12), e0243030 (2020)
7. Azimi, Y., Khoshrou, S.H., Osanloo, M.: Prediction of blast induced ground vibration (bigv) of quarry mining using hybrid genetic algorithm optimized artificial neural network. Measurement 147, 106874 (2019)
8. Ramchoun, H., Ghanou, Y., Ettaouil, M., Janati Idrissi, M.A.: Multilayer perceptron: Architecture optimization and training (2016)
9. Ramchoun, H., Idrissi, M.J., Ghanou, Y., Ettaouil, M.: Multilayer perceptron: Architecture optimization and training with mixed activation functions. In: Proceedings of the 2nd international Conference on Big Data, Cloud and Applications. pp. 1–6 (2017)
10. Moghaddam, M.A., Golmezergi, R., Kolahan, F.: Multi-variable measurements and optimization of gmaw parameters for api-x42 steel alloy using a hybrid bpnn–pso approach. Measurement 92, 279–287 (2016)
11. Socha, K., Blum, C.: An ant colony optimization algorithm for continuous optimization: application to feed-forward neural network training. Neural computing and applications 16, 235–247 (2007)
12. Castillo, P.A., Merelo, J.J., González, J., Rivas, V., Romero, G.: Sa-prop: Optimization of multilayer perceptron parameters using simulated annealing. In: Foundations and Tools for Neural Modeling: International Work-Conference on Artificial and Natural Neural Networks, IWANN'99 Alicante, Spain, June 2–4, 1999 Proceedings, Volume I 5. pp. 661–670. Springer (1999)
13. John, H.: Holland. adaptation in natural and artificial systems/john h. holland. Ann Arbor: University of Michigan Press (1975)
14. De Campos, L.M.L., de Oliveira, R.C.L., Roisenberg, M.: Optimization of neural networks through grammatical evolution and a genetic algorithm. Expert Systems with Applications 56, 368–384 (2016)
15. Hassibi, B., Stork, D.: Second order derivatives for network pruning: Optimal brain surgeon. Advances in neural information processing systems 5 (1992)
16. Attik, M., Bougrain, L., Alexandre, F.: Optimal brain surgeon variants for optimization. In: ECAI. vol. 16, p. 957 (2004)
17. Dasgupta, D., McGregor, D.R.: Designing application-specific neural networks using the structured genetic algorithm. In: [Proceedings] COGANN-92: International Workshop on Combinations of Genetic Algorithms and Neural Networks. pp. 87–96. IEEE (1992)

18. Sánchez, D., Melin, P.: Optimization of modular granular neural networks using hierarchical genetic algorithms for human recognition using the ear biometric measure. Engineering Applications of Artificial Intelligence 27, 41–56 (2014)
19. Deb, K., Jain, H.: An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, part i: solving problems with box constraints. IEEE transactions on evolutionary computation 18(4), 577–601 (2013)
20. Blank, J., Deb, K.: pymoo: Multi-objective optimization in python. IEEE Access 8, 89497–89509 (2020)