Predicting reinforcement fracture in concrete walls and identifying vulnerabilities

Juan Miguel Navarro Carranza*
Department of Civil and Environmental Engineering
Stanford University
jmnavarr@stanford.edu

1 Introduction

Buildings designed to withstand earthquake loadings aim to resist the generated forces by dissipating energy through undergoing inelastic behavior of specific structural parts. In concrete structures, the designated element to dissipate energy is reinforcing steel. These inelastic cycles degrade the fatigue life cycle of reinforcing steel and can cause premature fracture which deteriorates the strength of the buildings. Extracted reinforcing bars from structural walls that were part of a damaged building in the 2010 Chile earthquake were tested to evaluate their low-cycle fatigue life. It was found that they had a reduction of about 40% of their fatigue capacity compared to intact bars meaning that those bars would not have been able to undergo another similar earthquake due to the cumulative damage.

Modeling fracture in the steel reinforcement of a concrete wall building has proven to be a very challenging endeavour; however recently developed mathematical models permit the simulation of the response of reinforced concrete wall buildings under earthquake loading. Identifying the vulnerabilities of the current design building codes for reinforced concrete buildings is important to avoid premature fracture of reinforcing during an earthquake.

This project intends to leverage the power of ML and the recently developed physics-based models that are expensive to compute to create an efficient surrogate model and to determine the important features that make a wall vulnerable to fracture. These insights will enable us to provide better design provisions to make structural concrete more resilient to reinforcing fracture.

2 Dataset and features

Training a reasonable surrogate model can be challenging given the variety of inputs for the model, especially the loading coming from earthquakes. The main challenge to producing the data, is to sample wall configurations that represent the real distribution of actual wall designs that are built in real life. Random configuration could be worthless if there is not at least certain constrains when sampling the wall features. For example, producing a wall that contains less than 0.001% of steel or that uses irrational reinforcing bar diameters compared to the size of the wall could be as good as useless. An algorithm that includes several constrains to ensure reasonable values was implemented to sample the walls, these constrains include obtaining mean values from real wall databases and bounding the distributions from which the random values are generated to reasonable quantities. Additionally, building code constrains were considered.

The dataset is formed by a set of features of the walls including materials and dimensions. These features are numerical floating point variables which constitute the input variables for the physics-based model and the neural network. The output variable pertains to the categorical type of variables and corresponds to a Boolean indicating if there is fracture or not in the wall after the application of the loading.

*Ph.D in Structural engineering

The output of the dataset for this project is produced through simulations of the existing physics-based model that captures the behavior of a reinforcing concrete wall. The model receives as input material properties, specimen dimensions, and loading that were sampled as mentioned before. The output of this model is the response of the wall, from this response it is determined if there was fracture or not.

The number of input features considered in this study is 55 plus the Boolean output. In total, a set of 6336 simulations were performed, from these, there was 2832 specimens that fractured for a 45% of the complete dataset. In an effort to keep a balanced dataset not all the actual simulations were used which limited the size of the dataset to the number above given the computational cost of the physics based model. Examples of the configurations produced by the wall random sampling algorithm can be seen in Figure 1.

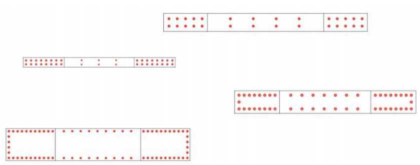

Figure 1: Sample configurations

3 Methods

The features for the model were analyzed before training and tuning of a fully connected neural network using Feature Selection methods to identify variables that can be important and that provide some insight of what it is relevant to the correct prediction of fracture. This feature evaluation was performed using supervised and unsupervised methods which difference is based on whether the output is accounted or not during its evaluation.

The type of problem that we have is a Classification Predictive Modeling problem since the model's goal is to map floating point numerical types of variables (features) to a categorical Boolean type of variable (output, i.e fracture or no fracture). Therefore methods that we have used to perform a Feature Selection analysis that fit our problem data include ANOVA-f Statistics and Mutual Information Statistics which combined with an algorithm such as Select K best can provide us with a selection of features that are more relevant to our problem.

The project intends to produce a surrogate model to substitute the physics-based model using Deep Neural Networks. The method chosen for the surrogate model was a Deep Neural Network with fully connected layers. In order to train our model, the neural network uses as the cost function a "Binary Cross Entropy" function given the nature of our data.

Methods of Explainable Machine Learning and Feature Importance were applied to the trained model. One of the methods uses SHAP (SHapley Additive exPlanations) [6] which uses a game theoretic approach that enable us to explain the the behavior of any machine learning model and its outputs. This method combines local explanations using Shapley values and optimal credit allocation.

4 Experiments/Results/Discussion

4.1 Feature analysis of dataset

The ANOVA-f Statistics and Mutual Information Statistics combined with the Select K best algorithm were developed to perform initial feature selection. The ANOVA-f Statistics method highlights some of the steel properties as the more relevant for the prediction of fracture as can be seen in Figure 2. The most important features according to this method include the ultimate strains of the steel which is the strain at which a bar monotonically loaded will achieve maximum stress. This result is not surprising, additionally ANOVA-f detects the strain of the three types of steel considered in the model, and as a matter of fact these types of steel are usually correlated.

The Mutual Information method returns very similar results as ANOVA-f prioritizing steel properties, however it also includes some properties related to the spacing of the transversal steel that prevents bars from buckling which accelerates fracture as can be sen in Figure 3.

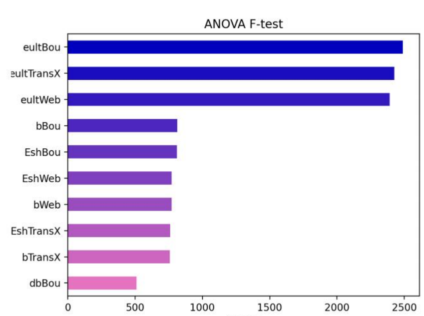

Figure 2: ANOVA f-test

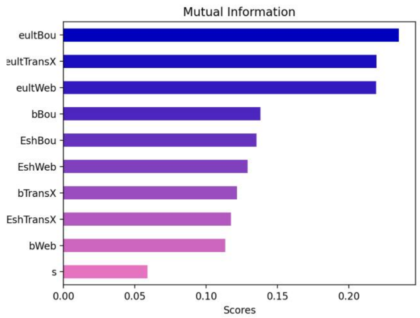

Figure 3: Mutual Information

4.2 Deep Neural Network Architecture

The architecture of the surrogate model is composed by an input of 55 features with 6336 samples followed by 30 fully connected layers with 30 neurons each. The hidden layers are followed by ReLU activation whereas the output layer except is a single neuron with Sigmoid activation to provide a probability that can be converted into a binary classification and feed the Binary Cross Entropy cost function. The model was built using the Tensorflow framework.

The optimizer used was ADAM with a learning rate of 0.0001. The final number of epochs used was 2500 with batch size of 80. Using 80% of the dataset for training and 20% for testing results in the Loss history shown in Figure 4 where the training loss is in light blue and the test loss is presented in fuchsia. As for the accuracy of the model, the results can be observed in Figure 5 with the same color code. The model had a 91% accuracy in the training set and 90.2% in the test set which denotes that the model was able to learn important features of the wall with little overfitting. Other architectures including Batch Normalization and Dropout were used, but did not provide any improvement. Architectures with more layers did not provide an improvement either and they even caused overfitting without accuracy improvement in the test set.

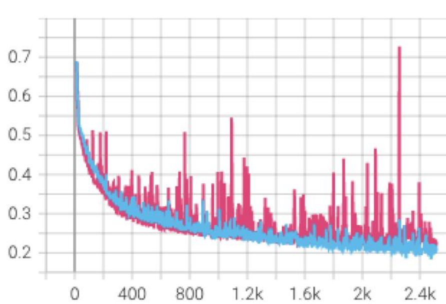

Figure 4: Loss history

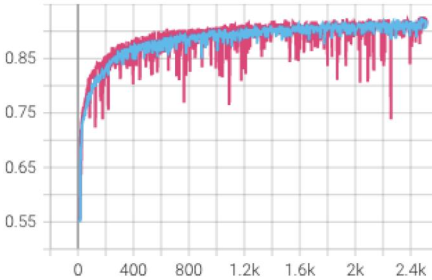

Figure 5: Accuracy history

4.3 Explaining the Deep Neural Network

The Deep Neural Network performs at a considerably good level of accuracy, however to understand why it takes such decisions the SHAP scores were obtained to try to explain how the neural network is behaving. Explaining how the surrogate neural network performs will provide insights about the actual phenomena happening in the the structural concrete walls and the causes of fracture in the reinforcing steel.

The SHAP scores were computed by integrating 1000 data points into an explainer function from the training set and calculating the scores for the complete test set. Each feature for each of the data points from the test sets receives a SHAP score according to its influence in the decision of the deep neural network. A summarized plot of the results can be observed in Figure 6. The SHAP scores reveal that the model prioritizes as well several of the steel properties, however it also includes dimensions of the wall. This set of features might seem random at first compared to the Feature Selection methods described before, but they are indeed quite reasonable.

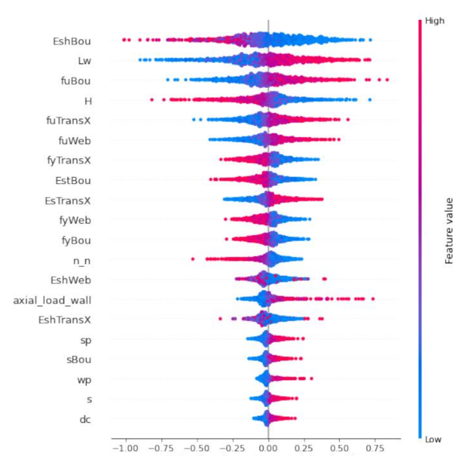

Figure 6: Overall behavior of NN

It is well known that the dimensions of the wall can affect the strain that the bars can experiment and this is well captured by the model as we can see in Figure 6 where the length of the wall (Lw) causes fracture when it has a large value which is indeed very logical because the larger the distance of the bar from the neutral axis of the cross section the larger the strain which is in turn a direct cause of fracture. In the plot the high values of the feature are expresed in fuchsia whereas blue colors are for lower values. The horizontal axis expresses the SHAP value and a positive value indicates that the feature is causing fracture, this means a fuchsia color to the right means a high value of that feature influences fracture according to the model. Each point in the plot is a SHAP score for each test point for each feature. The density expresses the Number of points wth a determined value for each feature. Therefore, more important features are on top.

Another insight is the fact that the model prioritizes overall as the most important feature the post yielding slope of the steel in the boundary element (EshBou) this indeed is a deep insight since this value is an important indicator of how resistant is a material or overall a structure after yielding. The model catches this value and takes it in consideration when taking decisions, this is a remarkable behavior. These insight could suggest that to make a wall resilient to fracture the steel material used in the bars should have a larger post yielding slope.

4.4 Explaining the Deep Neural Network for single specimens

So far the overall behavior of the neural network was described, but a local behavior can also be analyzed. This means that for an specific sample we can get the scores for each feature and how they affected the decision of the neural network. Let's look at the test set example of a fractured wall depicted in Figure 7 .

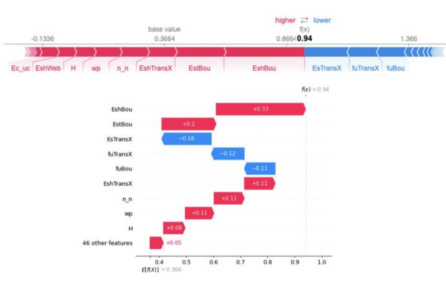

Figure 7: Example of fractured specimen

In the fractured example the neural network considers again that the relatively low value of the post yielding slope of the steel in the boundary element will cause fracture, this indeed was accurate since it predicted fracture with 94% of probability. In this figure the red values push the model towards considering fracture where as the blue ones try to push the model to consider that it will not fracture. The contribution of each feature ends up adding up to the model predicting fracture correctly.

Let's look now at an example from the test set where there was not fracture. In Figure 8 we can see how the value of the post yielding slope of the boundary steel dominates the decision of the model, but in this case a large value causes the model to predict correctly that there is not fracture in the specimen.

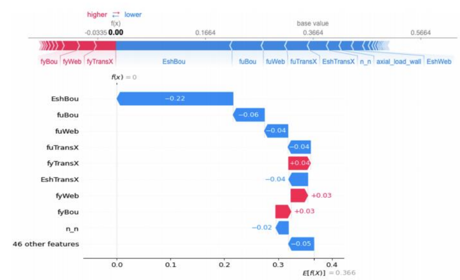

Figure 8: Example of not fractured specimen

5 Conclusion/Future work

The surrogate model developed using Deep Neural Networks was able to have a high accuracy predicting fracture, which is not a trivial task. The model was able to learn important features that reflect insights from the actual behavior of structural concrete walls and its components which indicates that the decisions are more than aleatory.

This was confirmed after applying explainable machine learning methods such as SHAP which helped us to identify the most important features causing fracture in the wall according to the surrogate model. Beyond the features that were considered important by the author before, the explainable ML methods revealed insights about other features that might be more relevant than expected, such as the slope of the steel after yielding.

These important features can be used to also improve the surrogate model by increasing targeted simulations for training and most importantly to detect vulnerabilities that should be treated when designing a reinforced concrete wall building. Explainable machine learning can give us insights to improve wall design to prevent fracture during earthquakes and prove to be powerful tools to help us understand physical phenomena.

Future work include increasing the training dataset to improve the model as well as to refine the constrains for the random sampling to improve the configurations of the walls to approach reality as much as possible. Additionally, extending the surrogate model to dynamic analysis that can help predict the exact time of fracture to try to predict the actual residual fatigue life of fractures and help improve its behavior and provide further insights.

6 Contributions

A surrogate model that is able to predict if a wall structure is going to have fracture giving its configuration was developed. Valuable insights from applying explainable machine learning to the current model were obtained relating to what features affect the most the fracture of reinforcing steel in structural concrete shear walls. These insights can help us define improvements while designing a structural concrete wall to make it resilient to fracture.

7 Literature

The ideas for the surrogate model and explainable methods were sparked after the following readings: [6], [8], [4], [2], [10], [3], [11], [9], [7], [1] and [5].

References

[1] Anupam Datta, Shayak Sen, and Yair Zick. "Algorithmic Transparency via Quantitative Input Influence: Theory and Experiments with Learning Systems". In: *Proceedings - 2016 IEEE Symposium on Security and Privacy, SP 2016* (Aug. 2016), pp. 598–617. DOI: 10.1109/SP.2016.42.

[2] Amirata Ghorbani et al. "Towards Automatic Concept-based Explanations". In: *Advances in Neural Information Processing Systems* 32 (Feb. 2019). ISSN: 10495258. DOI: 10.48550/arxiv.1902.03129. URL: https://arxiv.org/abs/1902.03129v3.

[3] *Interpretable Machine Learning*. URL: https://christophm.github.io/interpretable-ml-book/.

[4] Pang Wei Koh and Percy Liang. "Understanding Black-box Predictions via Influence Functions". In: *34th International Conference on Machine Learning, ICML 2017* 4 (Mar. 2017), pp. 2976–2987. DOI: 10.48550/arxiv.1703.04730. URL: https://arxiv.org/abs/1703.04730v3.

[5] Isaac Lage et al. "An Evaluation of the Human-Interpretability of Explanation". In: (Jan. 2019). DOI: 10.48550/arxiv.1902.00006. URL: https://arxiv.org/abs/1902.00006v2.

[6] Scott M Lundberg and Su-In Lee. "A Unified Approach to Interpreting Model Predictions". In: *Advances in Neural Information Processing Systems*. Ed. by I Guyon et al. Vol. 30. Curran Associates, Inc., 2017. URL: https://proceedings.neurips.cc/paper/2017/file/8a20a8621978632d76c43dfd28b67767-Paper.pdf.

[7] Scott M. Lundberg and Su In Lee. "A Unified Approach to Interpreting Model Predictions". In: *Advances in Neural Information Processing Systems* 2017-December (May 2017), pp. 4766–4775. ISSN: 10495258. DOI: 10.48550/arxiv.1705.07874. URL: https://arxiv.org/abs/1705.07874v2.

[8] Scott M. Lundberg et al. "Explainable machine-learning predictions for the prevention of hypoxaemia during surgery". In: *Nature Biomedical Engineering 2018 2:10* 2.10 (Oct. 2018), pp. 749–760. ISSN: 2157-846X. DOI: 10.1038/s41551-018-0304-0. URL: https://www.nature.com/articles/s41551-018-0304-0.

[9] Marco Tulio Ribeiro, Sameer Singh, and Carlos Guestrin. ""Why Should I Trust You?": Explaining the Predictions of Any Classifier". In: *NAACL-HLT 2016 - 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Proceedings of the Demonstrations Session* (Feb. 2016), pp. 97–101. DOI: 10.48550/arxiv.1602.04938. URL: https://arxiv.org/abs/1602.04938v3.

[10] Mukund Sundararajan, Ankur Taly, and Qiqi Yan. "Axiomatic Attribution for Deep Networks". In: *34th International Conference on Machine Learning, ICML 2017* 7 (Mar. 2017), pp. 5109–5118. DOI: 10.48550/arxiv.1703.01365. URL: https://arxiv.org/abs/1703.01365v2.

[11] Zifan Wang et al. “Interpreting Interpretations: Organizing Attribution Methods by Criteria”. In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops* 2020-June (Feb. 2020), pp. 48–55. ISSN: 21607516. DOI: 10.48550/arxiv.2002.07985. URL: https://arxiv.org/abs/2002.07985v2.