Development of a deep neural network with simulated ground motion synthetics for seismic hazard and risk assessment in the Hayward Fault region

The proposed research idea is based on the fact that recent advancement in machine learning in conjunction with high-performance computing is providing us with an unprecedented way of developing models for complex scientific and engineering applications. Earthquake science and engineering is a field where this robust computing framework would be extremely beneficial. In this area we lack adequate number of recorded data due to historical insufficient amount of sensor deployment. The advantages of developing neural network in this field will be as follows

- To be able to predict ground motion intensity on a regional scale without running astronomically expensive (both time and cost) physics-based full wave propagation code
- To be able to predict building damage and corresponding risk without resorting to nonlinear building response simulations for each of hundreds or more possible hazard scenarios
- As time and funding permits we would be able to incorporate infrastructure and utility risk on top of building damage risk

The exclusive advantage of predicting both earthquake hazard and risk using a machine learning platform would be both in terms of cost and time. Once the model is fully trained with reliable number of synthetics and benchmarked with the available recorded data we would be able to tell the expected level of hazard and risk at a particular location or on region for a given magnitude, hypocenter location, rupture scenario, distance and so on. This would obviate the execution of both costly geophysics and structural engineering codes for a random variation of a rupture scenario. Because for a given geologic fault we do not know the exact location of future hypocenter or the slip distribution and having the information of how a slight variation in ground motion model parameters results in consequent hazard and risk within minutes or hours would be of utmost interest to scientists and engineers. With existing methodologies what would take years, with this methodology we would be able to accomplish that in days. All we need is to train the model with sufficient amount of data and it will be as accurate as with the fidelity of the training data we feed into the model. Also, we can always improve our model based on new data we gather and feed into it. With our current ECP project we are at the stage of simulating many hazard scenarios with engineering frequency of interest and greater accuracy for the Bay Area. I would like to leverage this existing capability to develop the proposed endeavor.