DIT866 Writing assigment 1

Jonatan Hellgren

February 2021

1 Introduction

I have coosen to read and discuss the paper written by: DeVries et al., *Deep Learning of Aftershock Patterns Following Large Earthquakes.* from Nature 560, 632-634 (2018)

2 Discussion

The goal of the project written about in the paper is to model the spatial distribution of aftershocks after mayor earthquakes. Why such a model has a value isn't explicitly mentioned in the paper, however some obvious reasons why this is a useful thing to have is that it has the potential to save lifes and money to be able to forecast where potential aftershocks are likely to happenkk.

How they modeled the spatial area in this task was by dividing an area spanned 100km horizontally and 50km verically from each mainshock rupture plane in to 5km $\times$ 5km $\times$ 5km cubes. By utalizing this dicritization of the spatial area they where able to do solve the task using a binary classification. Where the two options where that the cube had either had an aftershock between one secound and one year efter the inital earthquake (represented as a 1) or that there hadn't occured a aftershock in that cell (represented as a 0).

The data they used was the stress-change that occured in the centroid of a cube after a mainshock and wheter or not an aftershock had occured in that cell after an mainshock. So basically they used the stress-change as the dependent variable and if a aftershock had occured as a dependent variable. The stress-change data where calculated using data from the SRCMOD online database of fininte-fault rupture models and the aftershock data was collected from the International Seismological Center. They have collected the data from 118 mainshocks and 162,741 aftershocks. For training the model they used 75% of the data and the remaining 25% was used for testing and evaluation the model. They also mention that they downsampled cubes without aftershocks during training.

To perform the binary classification they used a deep neural network consisting of six hidden layers with each of them consisting of 50 neurons each. The activation function used in these layers where a hyperbolic tanget function. For the last layer they only used one neuron as output, which is the standard for doing binary classification. The network was trained using Theano, an adaptive learning rate and optimization method, and for the cost function they used binary cross-entropy. The coice of this method wasn't mentioned, neither if they had tried any other similar models, neither did they mention if there where an drawbacks or potential error when going with this approach.

For evaluating the quality of the model they used the metric AUC, which measures the area under a ROC curve. A ROC curve i drawn by plotting the true positive rate and the true negative rate for different thresholds. This is a measurment that is defined in the interval $[0, 1]$ where a higher score is better and a value of one means that the classifier has 100% true positive rate and 0% false positive rate for every threshold.

The models performance was compared to a baseline that was measured using the a standard formula for these sorts of tasks called the Coulomb failure stress change. The AUC value their neural net model acheived was 0.849, compared to the score of 0.583 that the Coulomb failure stress change score this is quite an increase in predictability.

The forcasting shadow generated by the network closely resembeled other already existing physical properties, which could explain a majority of the varaince in the model. The paper then goes on to say that "These results highlight how deep-learning approaches can lead to improved aftershock forecasts and provide physical insights into the mechanisms of earthquake triggering", which is a good conclusion about the paper.

My own personal opinion about this project is that it is interesting to see how neural networks are able to find paricular patterns i things that we humans have a hard time picking up, primarily because our limitation of being able to process that amount of data. Even though the results sometimes may be informative it isn't always that way, deep neural networks are ofter refered to as a black box since it is pretty much impossible to interpret how they make there choices. This could lead to problems in this domain if one would become to cnfident in the predicitions of the network and migh end up harming themself or other people in the case of an earthquake. Another thing I feelt that the paper missed was a discussion on what other machine learning models that might be adequite for this task, or if there is any.