Earthquake Rupture Dynamics

using Machine Learning

Submitted by
Abhay Mishra (181CV201)
Sathvika.B.Mahesh (181CV141)
Saurabh Tiwari (181CV241)

Objective:

Dynamic rupture propagation is uncertain due to the factors that consider the working mechanism of faults, slip, strike, stress conditions, and friction of the fault. Using machine learning techniques, we predict if a rupture can break a fault.

We will be using two models namely Artificial Neural Networks and Random Forest Classifiers to train the data and taking out the results based on various metrics. Both the Algorithms back their roots from the field of Machine learning and Deep learning respectively. Their robust nature makes their use meaningful over here.

MOTIVATION:

Dynamic rupture propagation is a challenging problem due to uncertainty regarding the underlying physics of earthquake slip, and the stress conditions and frictional properties of fault are not well constrained.

These unknown initial stresses and friction combine with fault geometry to control the rupture process and determine the dynamics of slip and the resulting ground motions.

However, because the earthquake rupture problem is highly nonlinear, determining parameter values is often done by making simplifying assumptions combined with trial and error, which computationally and numerically expensive.

To improve our ability to determine reasonable friction and stress parameters, we use machine learning methods to develop models to predict if rupture can break through a fault with geometric heterogeneities.

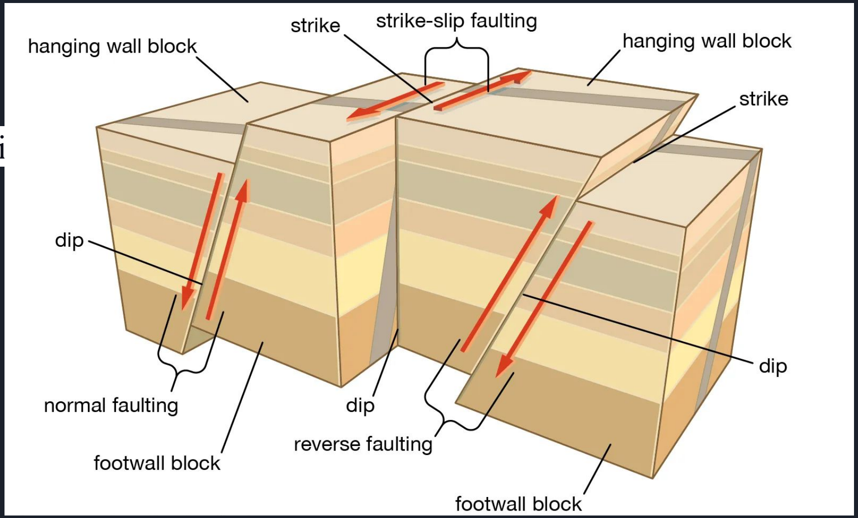
Understanding Earthquakes:
Important terminologies : Fault, Strike, Rupture, Barrier

Faults: are blocks of the earth's crust meeting
Together

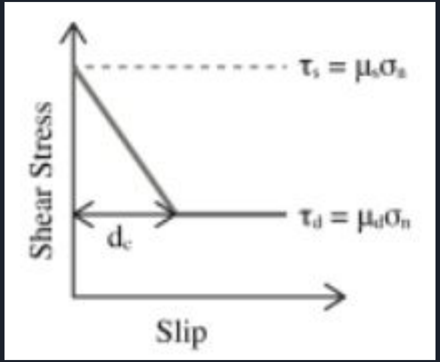
Strike: is the trend or bearing, relative to north, of the li
surface

Dip: is the angle at which a planar geologic
surface (for example, a fault) is inclined
from the horizontal.

Slip: is the relative displacement of formerly
adjacent points on opposite sides of a fault,
measured on the fault surface.



When does a fault begins to slip?

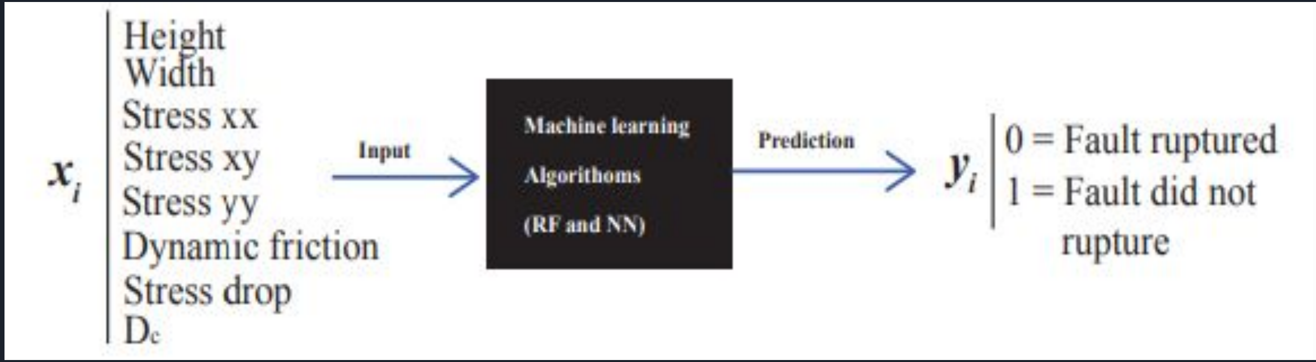


Slip-weakening friction for an earthquake fault. The fault begins to slip when the shear stress reaches or exceeds the peak strength τs. Over a critical slip distance dc, τs decreases linearly to a constant dynamic sliding friction τd. The shear strength is linearly proportional to the (possibly time-varying) normal stress σn, and the friction coefficient varies with slip between μs and μd

Methodology:

Using two machine learning algorithms

1. Random forest
2. Artificial Neural Network



Data preprocessing and Rupture Models:

The fault in the simulation is planar, with a Gaussian geometric heterogeneity at the center of the fault.

The rupture on the fault is governed by the linear slip-weakening law.

The fault starts to break when the
shear stress (τ) exceeds the peak strength $\tau_s=\mu_s\sigma_n$, where μ_s and σ_n are the static friction coefficient and normal stress, respectively.

The simulation dynamically determines fault slip based on the initial stress conditions, the elastodynamic wave equations, and frictional failure on that fault.

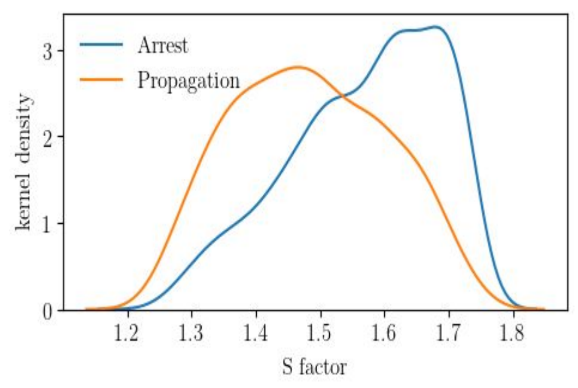
A dynamic rupture model solves the elastodynamic wave equation coupled to a friction law describing the failure process.

Data has been generated using fdfault (Daub, 2017) a finite difference code for numerical simulation of elastodynamic fracture and friction problems

Parameters:

Features to be used in dataframe :

1. Sxx and Syy normal stress
2. Sxy shear stress
3. Dynamic friction coefficient (mud)
4. Dc - critical slip distance
5. Friction drop - (static - dynamic friction coefficient)
6. Height and width of fault, others will be derived from them as S factor and mus.



- S-factor: The S-ratio is calculated as $S=(\tau_s-\tau_i)/(\tau_i-\tau_d)$, where τ_i is the initial shear stress, and τ_s and τ_d are the peak and residual shear stresses, respectively
- For this particular problem, however, we find that neither metric can reliably predict if the rupture can break the barrier for a given set of parameters.
- We find that the S-ratio does offer some utility as a discriminant for rupture, but only in the sense that large values of S are more likely to arrest.
- Smaller values of the S-ratio traditionally indicate a rupture that is closer to failure, but we find that for this particular problem small S-ratios still arrest in many situations.

1	Height	Width	Sxx	Syy	Sxy	Sdrop	Mud	Dc	Label
2	8.45E-02	1.02E+00	-2.36E+01	7.93E+00	-2.45E+01	2.54E-01	4.89E-01	3.88E-01	0.00E+00
3	1.34E-01	1.91E+00	-7.01E+01	3.51E+01	-7.45E+01	4.55E-01	3.26E-01	3.96E-01	0.00E+00
4	4.28E-02	1.66E+00	-7.38E+01	4.23E+01	-8.87E+01	4.43E-01	2.69E-01	4.60E-01	1.00E+00
5	4.61E-02	1.02E+00	-1.51E+02	4.88E+01	-1.30E+02	3.65E-01	2.73E-01	3.52E-01	0.00E+00
6	1.39E-01	1.73E+00	-6.90E+01	3.35E+01	-6.11E+01	5.36E-01	2.01E-01	4.15E-01	0.00E+00
7	9.46E-02	1.52E+00	-9.76E+01	6.44E+01	-1.19E+02	5.28E-01	2.59E-01	4.05E-01	0.00E+00
8	4.91E-02	1.56E+00	-2.84E+01	5.13E+00	-2.31E+01	2.02E-01	4.74E-01	3.64E-01	0.00E+00
9	1.05E-01	1.49E+00	-8.63E+00	5.39E+00	-1.01E+01	5.10E-01	2.32E-01	4.15E-01	0.00E+00
10	1.25E-01	1.52E+00	-2.60E+01	1.16E+01	-3.17E+01	3.51E-01	2.54E-01	3.91E-01	0.00E+00
11	1.06E-01	1.82E+00	-1.14E+02	2.96E+01	-1.06E+02	2.44E-01	2.71E-01	3.73E-01	1.00E+00
12	1.53E-01	1.92E+00	-1.44E+02	4.10E+01	-1.15E+02	3.34E-01	2.73E-01	3.36E-01	1.00E+00
13	9.92E-02	1.79E+00	-1.85E+02	3.55E+01	-1.59E+02	2.08E-01	4.78E-01	3.38E-01	1.00E+00
14	1.01E-01	1.35E+00	-2.26E+01	7.69E+00	-2.62E+01	2.84E-01	2.44E-01	3.62E-01	0.00E+00
15	1.09E-01	1.50E+00	-4.42E+01	2.18E+01	-5.34E+01	4.00E-01	3.24E-01	4.87E-01	0.00E+00
16	5.84E-02	1.67E+00	-3.08E+01	2.14E+01	-3.99E+01	5.28E-01	2.12E-01	4.39E-01	0.00E+00
17	4.92E-02	2.00E+00	-9.31E+01	4.08E+01	-7.79E+01	5.06E-01	2.41E-01	4.43E-01	0.00E+00
18	7.91E-03	1.36E+00	-1.56E+02	9.49E+01	-1.53E+02	5.93E-01	2.04E-01	4.65E-01	1.00E+00
19	1.29E-01	1.48E+00	-1.95E+01	9.13E+00	-1.58E+01	5.62E-01	2.18E-01	3.56E-01	0.00E+00
20	5.36E-02	1.88E+00	-3.87E+01	1.47E+01	-4.73E+01	2.61E-01	5.11E-01	4.31E-01	0.00E+00
21	4.54E-02	1.26E+00	-8.53E+01	3.15E+01	-1.11E+02	2.17E-01	5.30E-01	3.09E-01	1.00E+00

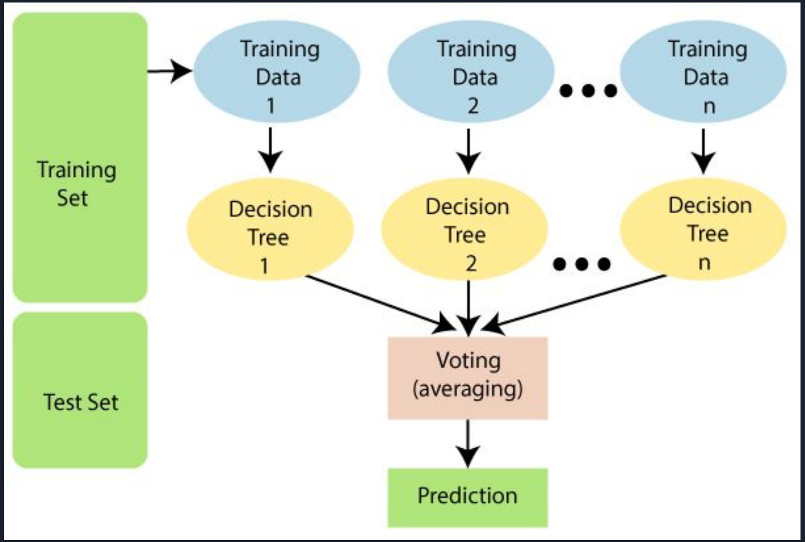
Training Dataset

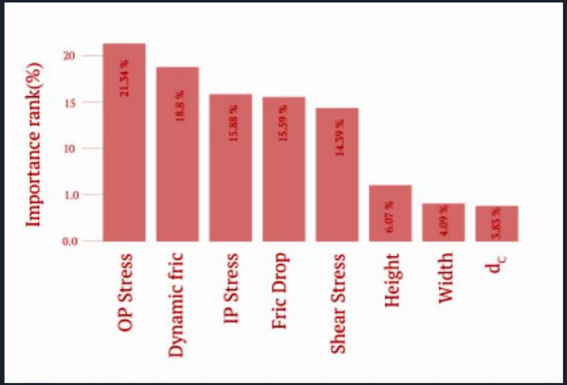
Testing Dataset

1	Height	Width	Sxx	Syy	Sxy	Sdrop	Mud	Dc
2	9.511999999999999622e-03	1.947726999999999986e+00	-6.513609700000000657e+01	2.318770800000000065e+01	-6.284791200000000089e+01	3.427100000000000146e-01	3.591750000000000220e-01	3.643100000000000227e-01
3	2.438900000000000096e-02	1.035033000000000092e+00	-1.067157490000000024e+02	2.342877199999999860e+01	-9.387410099999999602e+01	2.109979999999999911e-01	2.869180000000000064e-01	3.160649999999999848e-01
4	3.105499999999999927e-02	1.306248999999999993e+00	-1.417709230000000105e+02	5.133925800000000095e+01	-1.167437000000000040e+02	4.022310000000000052e-01	2.980869999999999909e-01	4.278089999999999948e-01
5	1.812729999999999897e-01	1.965154000000000067e+00	-6.409806799999999782e+01	1.682337800000000172e+01	-5.223822700000000196e+01	2.915530000000000066e-01	4.535009999999999875e-01	3.639990000000000170e-01
6	3.657000000000000094e-03	1.175945999999999936e+00	-9.095431000000000665e+01	2.892623499999999837e+01	-9.691538300000000561e+01	2.774249999999999772e-01	4.223049999999999859e-01	4.114380000000000259e-01
7	1.487600000000000033e-02	1.149683000000000010e+00	-1.121497889999999984e+02	5.593223799999999812e+01	-1.398614650000000097e+02	3.861140000000000128e-01	3.861760000000000193e-01	4.151529999999999943e-01
8	3.588600000000000123e-02	1.631131000000000109e+00	-1.424465340000000140e+02	7.743747899999999618e+01	-1.486255400000000009e+02	4.847839999999999927e-01	3.064640000000000142e-01	3.136840000000000184e-01
9	1.480300000000000019e-02	1.366522000000000014e+00	-1.399372250000000122e+02	7.000553600000000642e+01	-1.151516770000000065e+02	5.947829999999999506e-01	2.039420000000000122e-01	4.210610000000000186e-01
10	2.619200000000000000e-02	1.450693999999999928e+00	-1.423017069999999933e+02	5.058640100000000217e+01	-1.153897670000000062e+02	4.231730000000000214e-01	2.166319999999999912e-01	3.991999999999999993e-01
11	1.160600000000000000e-02	1.219273000000000051e+00	-1.157738020000000034e+02	4.699078300000000041e+01	-1.313502599999999916e+02	3.463499999999999912e-01	2.233570000000000000e-01	3.544229999999999881e-01
12	7.078299999999999870e-02	1.033647999999999900e+00	-1.182996550000000013e+02	6.155127499999999685e+01	-1.091290139999999980e+02	5.563160000000000327e-01	2.622290000000000134e-01	4.259189999999999920e-01
13	8.419000000000000095e-02	1.018283000000000049e+00	-1.719355590000000120e+02	5.598272899999999908e+01	-1.539061700000000030e+02	3.150200000000000222e-01	3.261049999999999782e-01	4.545690000000000008e-01
14	1.277439999999999964e-01	1.363034999999999997e+00	-9.832946099999999490e+01	3.541948000000000008e+01	-1.255623750000000030e+02	2.615419999999999967e-01	5.308230000000000448e-01	3.914320000000000022e-01
15	9.355099999999999527e-02	1.658643000000000089e+00	-7.925943700000000547e+01	3.440039600000000064e+01	-9.505259900000000073e+01	3.554930000000000034e-01	3.093489999999999851e-01	3.618529999999999802e-01
16	7.421400000000000219e-02	1.352573000000000025e+00	-1.352602149999999881e+02	9.761820699999999817e+01	-1.585147880000000100e+02	5.984519999999999840e-01	2.005099999999999938e-01	4.345300000000000273e-01
17	9.702800000000000313e-02	1.105812999999999935e+00	-5.537648999999999688e+01	3.646455300000000221e+01	-6.131797699999999907e+01	5.671279999999999655e-01	2.017299999999999927e-01	3.726940000000000253e-01
18	1.815660000000000052e-01	1.831625000000000059e+00	-1.233535810000000055e+02	7.067404200000000003e+01	-1.594594089999999937e+02	4.230809999999999849e-01	3.341250000000000053e-01	4.419359999999999955e-01
19	3.872500000000000248e-02	1.253133000000000052e+00	-1.022474940000000032e+02	4.062814600000000098e+01	-1.318216020000000128e+02	2.799690000000000234e-01	3.736260000000000137e-01	3.915170000000000039e-01
20	1.362520000000000120e-01	1.862122000000000055e+00	-1.190633710000000036e+02	5.015776300000000276e+01	-1.010882689999999968e+02	4.746299999999999963e-01	3.064479999999999982e-01	3.350009999999999932e-01
21	1.169789999999999996e-01	1.179024000000000072e+00	-1.387522800000000132e+02	9.138170499999999663e+01	-1.515684379999999862e+02	5.974739999999999496e-01	2.011139999999999872e-01	4.859530000000000238e-01
22	1.395599999999999966e-02	1.445240999999999998e+00	-6.261948999999999899e+01	1.572529499999999913e+01	-5.732236300000000284e+01	2.429759999999999975e-01	2.776270000000000127e-01	4.352559999999999762e-01
23	7.750899999999999457e-02	1.415751999999999899e+00	-4.627779999999999916e+01	2.455976400000000126e+01	-4.587442099999999812e+01	5.014929999999999666e-01	2.615009999999999835e-01	3.437850000000000072e-01
24	5.400699999999999945e-02	1.971832000000000029e+00	-7.739274100000000089e+01	4.136100400000000121e+01	-7.098640299999999570e+01	5.535750000000000393e-01	2.410249999999999893e-01	4.446550000000000225e-01
25	2.127100000000000171e-02	1.233046000000000086e+00	-1.093945329999999956e+02	2.864752599999999916e+01	-8.902151899999999785e+01	2.620060000000000167e-01	5.327429999999999666e-01	4.346570000000000156e-01

Random forest:

1) Contains a number of decision trees on various subsets of a given dataset
2) Takes the average to enhance the predicted accuracy of that dataset
3) Rather than depending on a single decision tree, the random forest collects the forecast's decision from each and every tree
4) Predicts the final output based on the majority votes of predictions





Importance feature

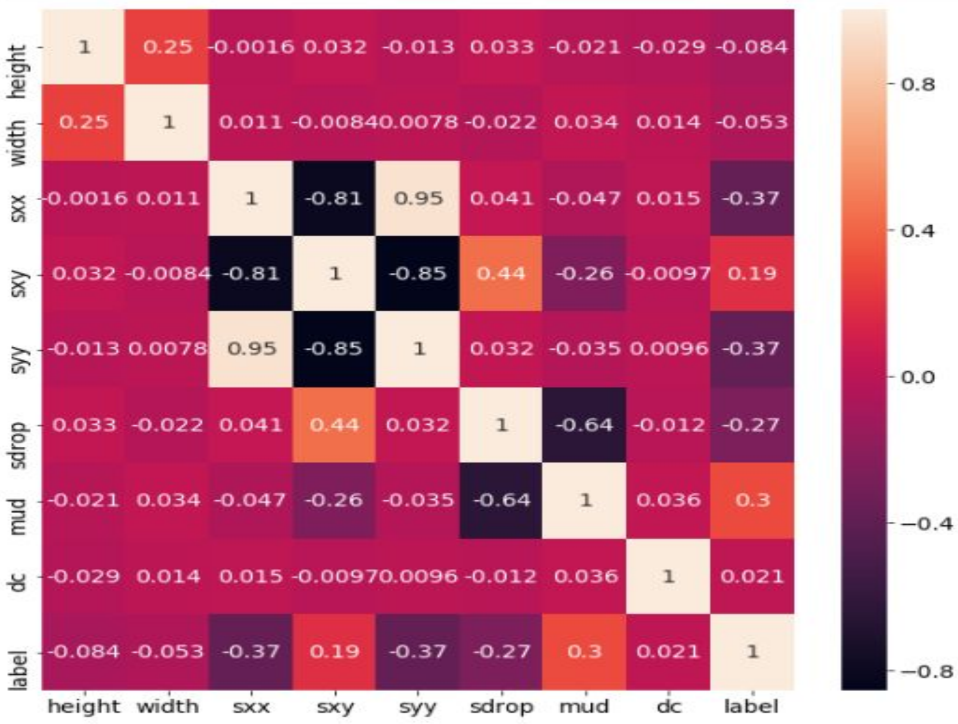
Classification Result

The accuracy of the model on 400 testing data set is **81.25%**

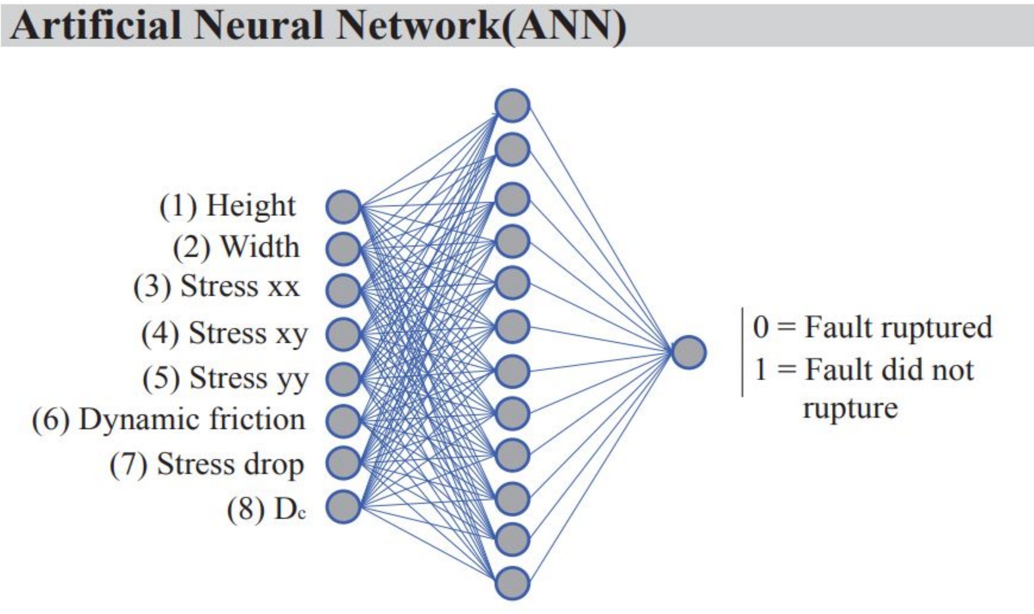
Class	Precision	True Positive Rate(Recall)	F1-score (Harmonic mean of precision and TPR)	No. of data
No Rupture (0)	0.91	0.80	0.85	272
Rupured (1)	0.66	0.84	0.74	128
Average/Total	0.83	0.81	0.82	400

HeatMaps

```python
plt.figure(figsize=(10, 10))
plt.rcParams.update({'font.size': 14})
sns.heatmap(df_train.corr(), annot = True)
plt.savefig('corr_heat_map.eps')
```



Artificial Neural Network:

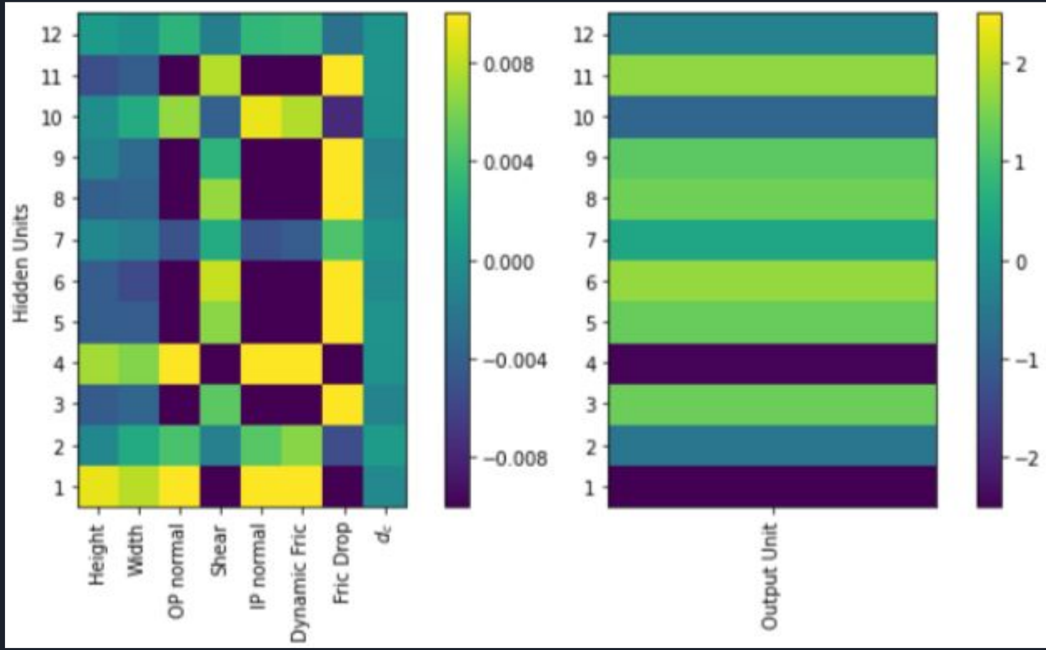


Classification Result

The accuracy of the model on 400 testing data set is **81 %**

Class	Precision	True Positive Rate(Recall)	F1-score (Harmonic mean of precision and TPR)	No. of data
No Rupture (0)	0.91	0.80	0.85	272
Rupured (1)	0.66	0.83	0.74	128
Average/Total	0.83	0.81	0.81	400

Rupture Models Parameters:



1) *Twelve nodes, the network has hidden layer*
2) *Weights that map the inputs to the hidden units are shown in the left panel.*
3) *The horizontal scale represents the eight input parameters*
4) *The vertical scale represents the twelve hidden units*
5) *Each row's colors show how the parameters are combined to create each hidden unit*
6) *The weights that are applied to each concealed unit to generate the single output unit are shown in the right panel*
7) *A large negative weight of the output unit indicates that the feature combination is predictive of arrest*
8) *A big positive value of the output unit shows that the combination is predictive of propagation*

Conclusion:

- Both models exhibit roughly 84 percent recall on the positive class (rupture propagation), indicating that they both understood much of the underlying complicated data patterns that cause rupture propagation.

- If a fault has higher out-of-plane normal and shear stress as well as low static and dynamic friction, an earthquake rupture is more likely to occur, and our model allows us to quantify this.

- The models are extremely computationally efficient. The models can produce a forecast in a fraction of a second after the training simulations are computed and the machine learning algorithms are taught.

Conclusion:

- Random forest and ANN models predict rupture propagation with greater than 81 percent accuracy.
- The underlying complicated data patterns that affect the mechanics of earthquake rupture propagation can be identified through machine learning.
- The method can also be used to solve other complex rupture problems like branching faults, fault stepovers, and other complex heterogeneities where the physics of earthquake rupture propagation aren't well known.

Code:

https://colab.research.google.com/drive/19mGSUsMZc6-fg6SVh6EA16DPzPtnC4Ew#scrollTo=vV3ehNeg4GIT

https://colab.research.google.com/drive/13Yo7r0JQY3SK_5fhIYzsZ2kSUIlL7rnd-

Outcome:

```
Accuracy of the model: 0.82

Classification report:
              precision    recall  f1-score   support

         0.0       0.86      0.87      0.87       669
         1.0       0.73      0.72      0.73       331

    accuracy                           0.82      1000
   macro avg       0.80      0.79      0.80      1000
weighted avg       0.82      0.82      0.82      1000


Confusion matrix:
[[582  87]
 [ 93 238]]
```

ANN test data prediction

```
The accuracy of the model is 0.8125
# Classification report
             precision    recall  f1-score   support


        0.0       0.90      0.81      0.85       272
        1.0       0.67      0.81      0.73       128

avg / total       0.83      0.81      0.82       400
```

Confusion matrix

```
confusion_matrix(ypred, ytest)
```

```
array([[221,  24],
       [ 51, 104]])
```

Random forest test Data Prediction

Thank you