



# Expanding Machine Learning to Formation Evaluation

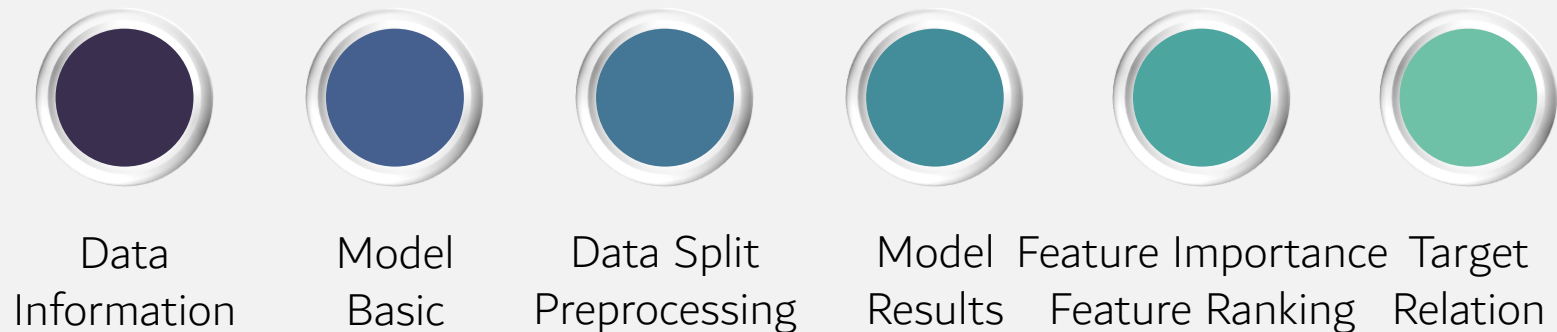
Neha Patel

Student, Flatiron School

# Objectives

- To propose a model that would accurately determine the type of facies based on the log readings speedily.
- To accurately determine the feature importances based of the machine learning model to aid in logging method selection
- Analyze the relationship between different log readings, and geological locations with facies classification

## Presentation Flow



# Data Information

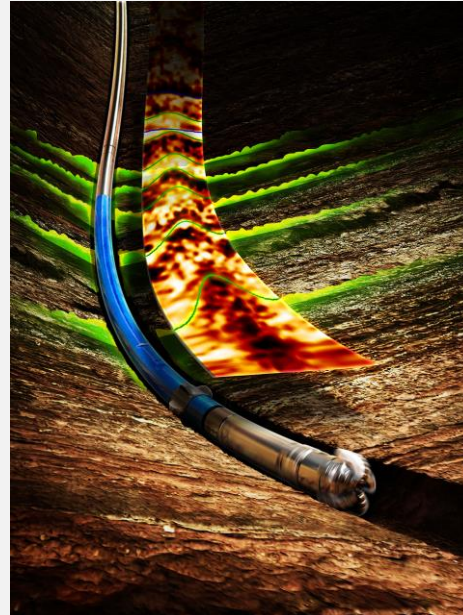


Data  
Information



## T A R G E T

- 1 : Non-Marine Sandstone
- 2: Non-Marine Coarse Siltstone
- 3 : Non-Marine Fine Siltstone
- 4 : Marine siltstone and shale
- 5 : Mudstone (Limestone)
- 6 : Wackestone (Limestone)
- 7 : Dolomite
- 8 : Packstone - Grainstone (Limestone)
- 9 : Phylloid-algal bafflestone (Limestone)



## F E A T U R E S

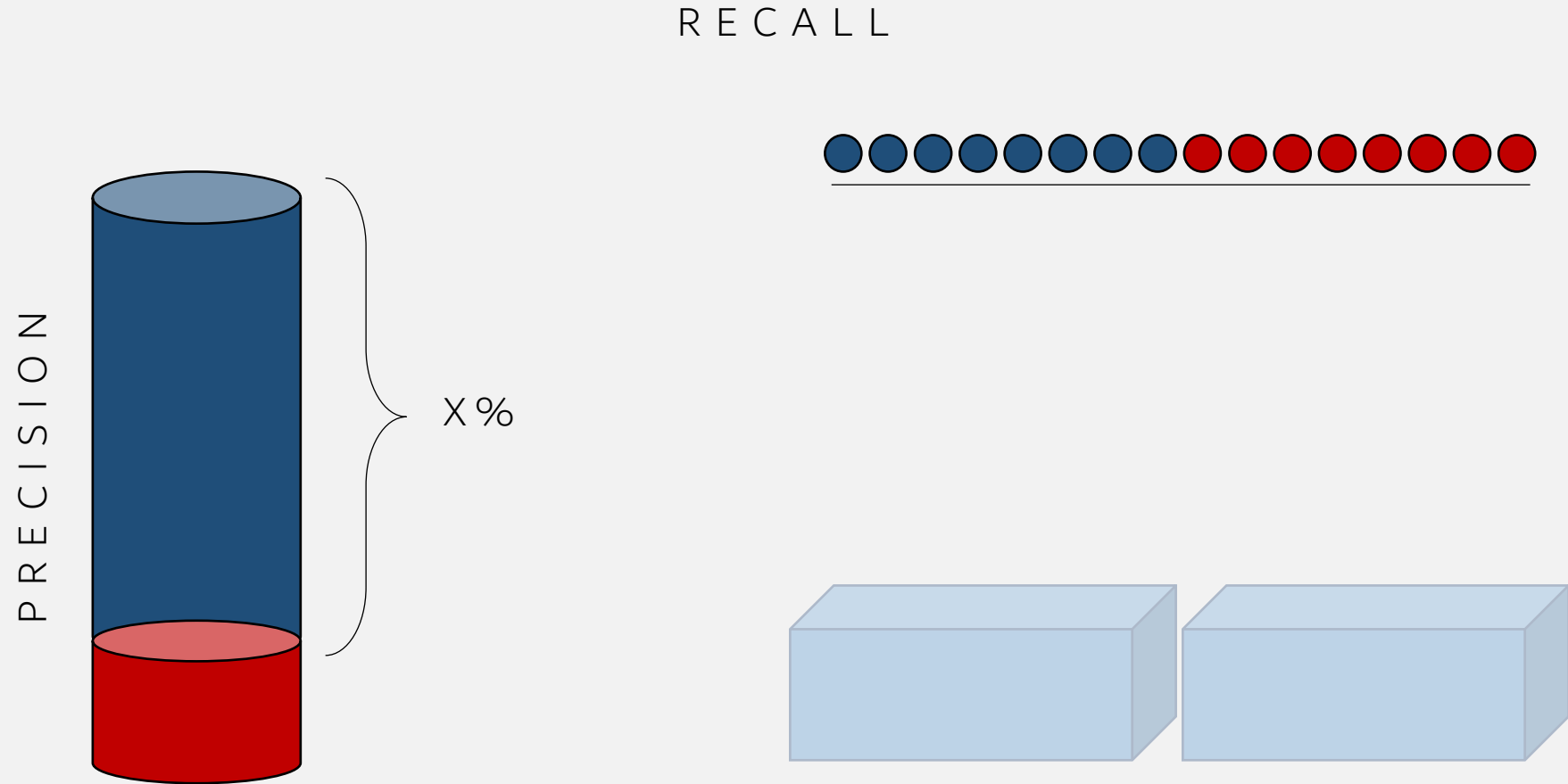
### Log Readings:

- Gamma Ray (GR)
- Resistivity (ILD\_log10)
- Photoelectric Effect (PE)
- Neutron-Density Porosity Difference (DeltaPHI)
- Neutron-Density Porosity (PHID)

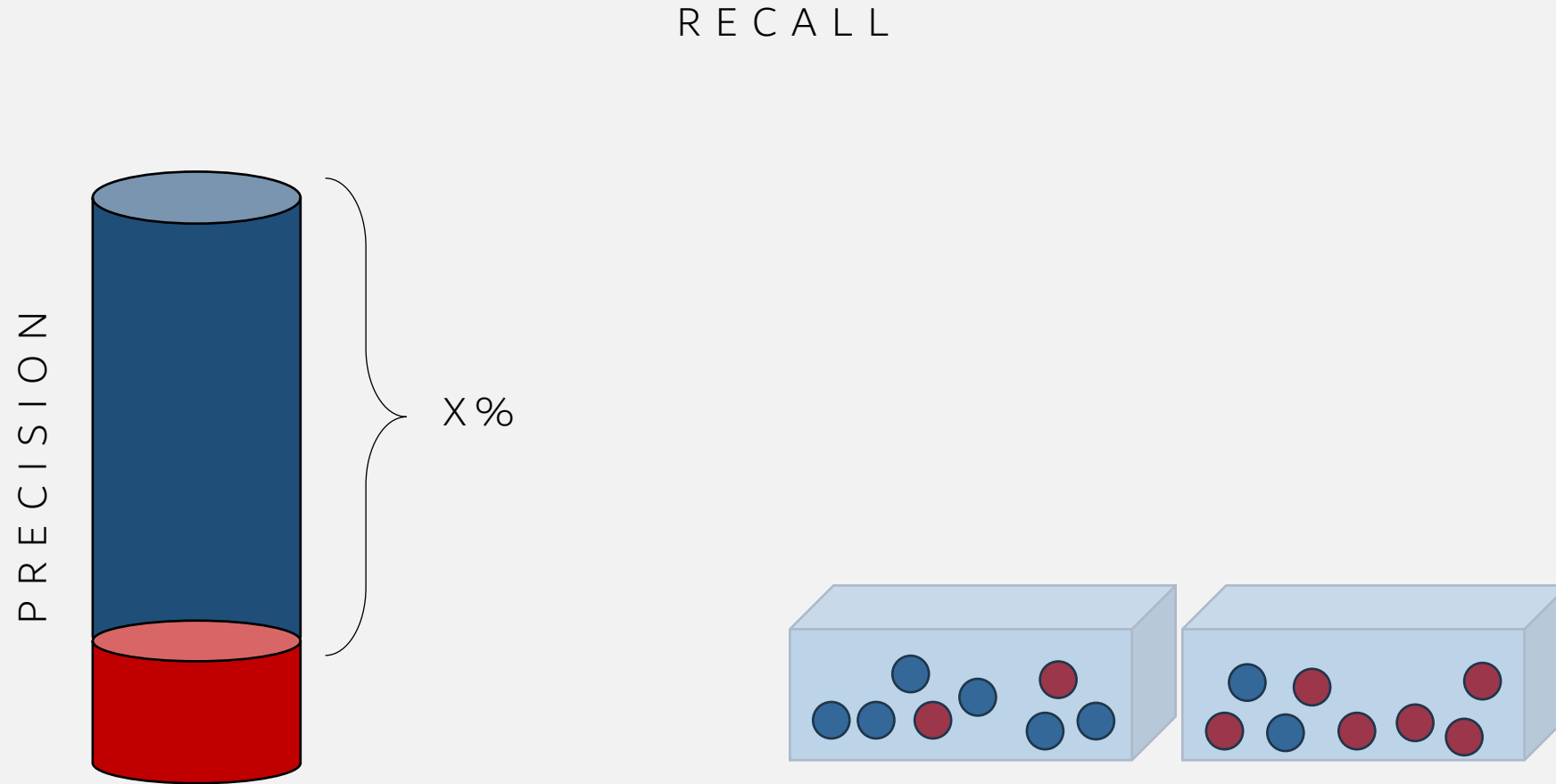
### Positional/Geological

Depth, Nonmarine-Marine Indicator (NM\_M), relative position, Formation and well names

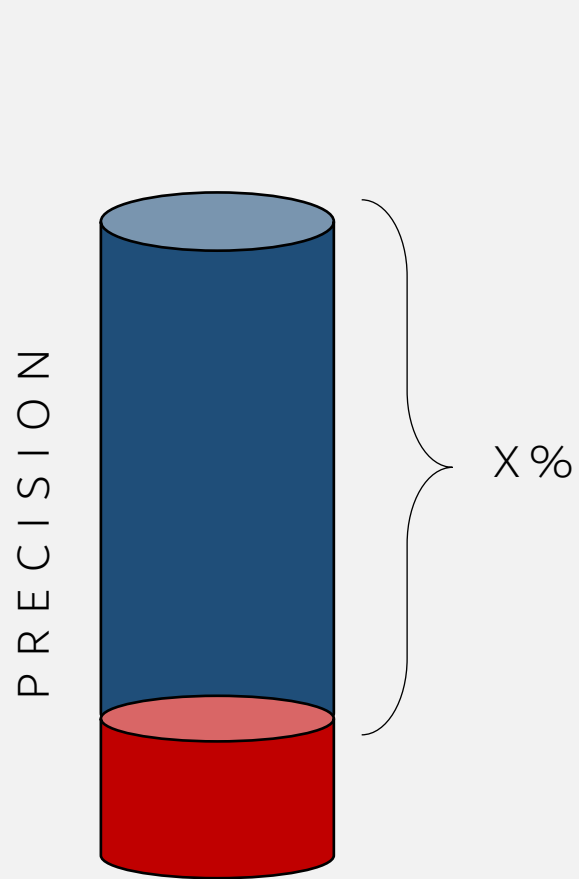
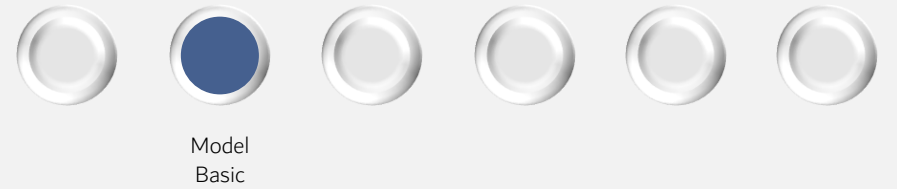
# Model Basics – F1 Score



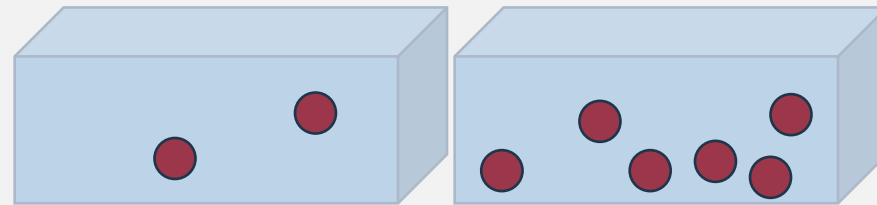
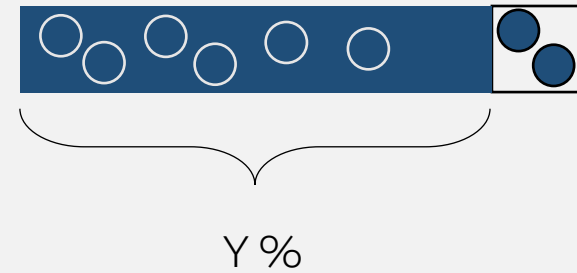
# Model Basics – F1 Score



# Model Basics – F1 Score



R E C A L L

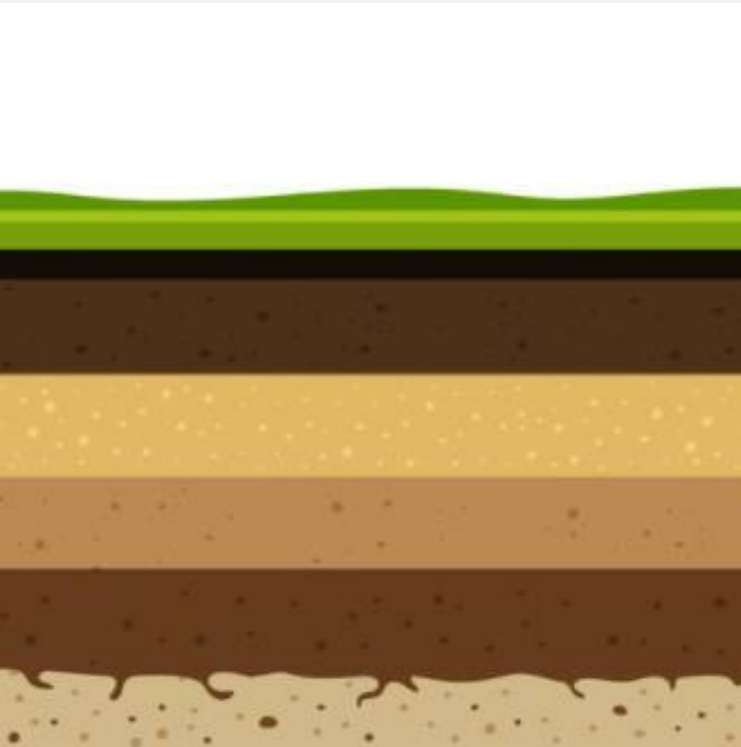




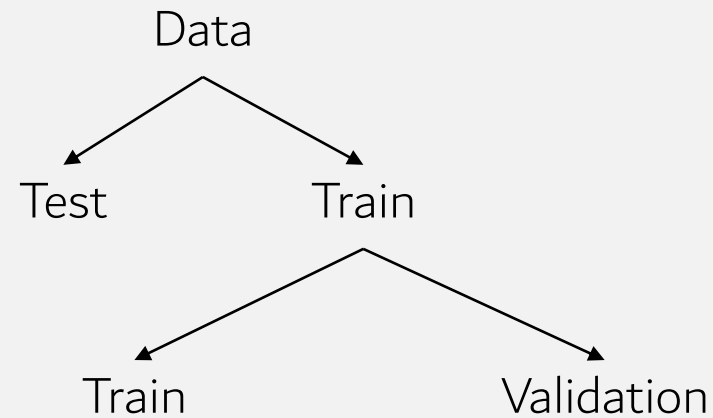


Data Split  
Preprocessing

# Data Split And Preprocessing



## DATA SPLITTING



## PREPROCESSING

Filling Missing Values:  
Using Distances

Scaling:  
Using Median

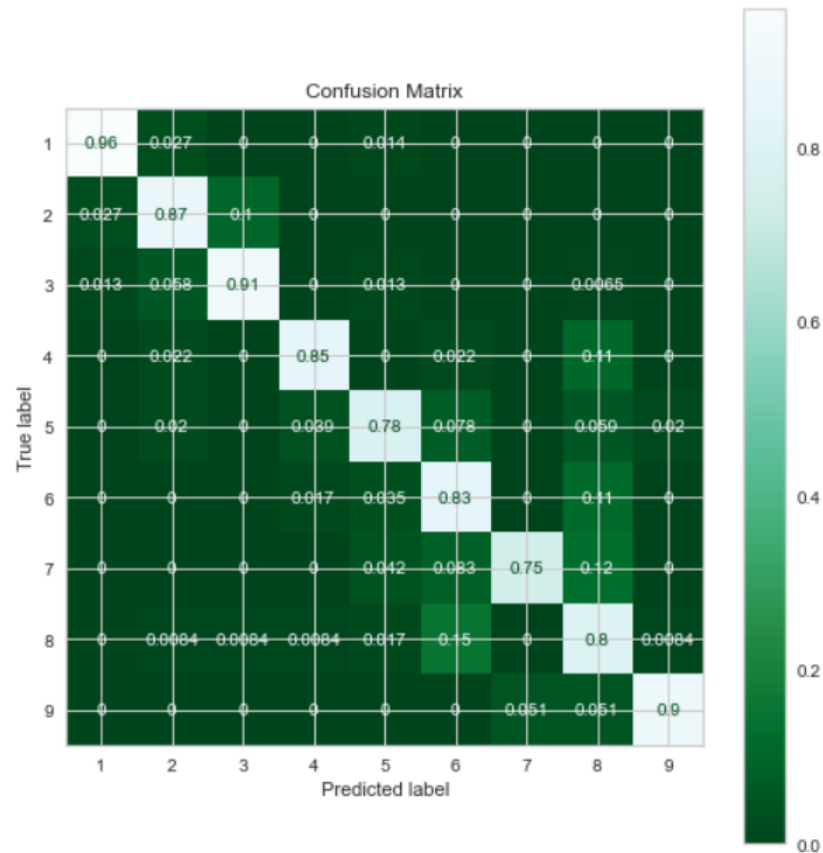


Model  
Results

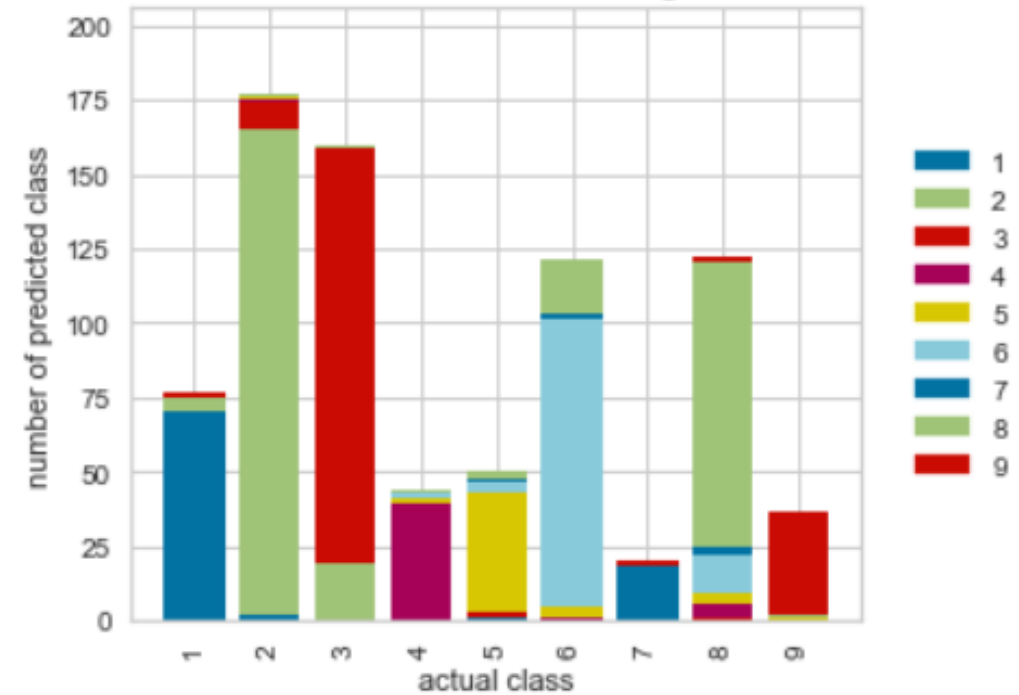
# Results on the Test Data

Time taken: 11.1s

## Visual Evaluation



## Class Prediction Error for StackingClassifier



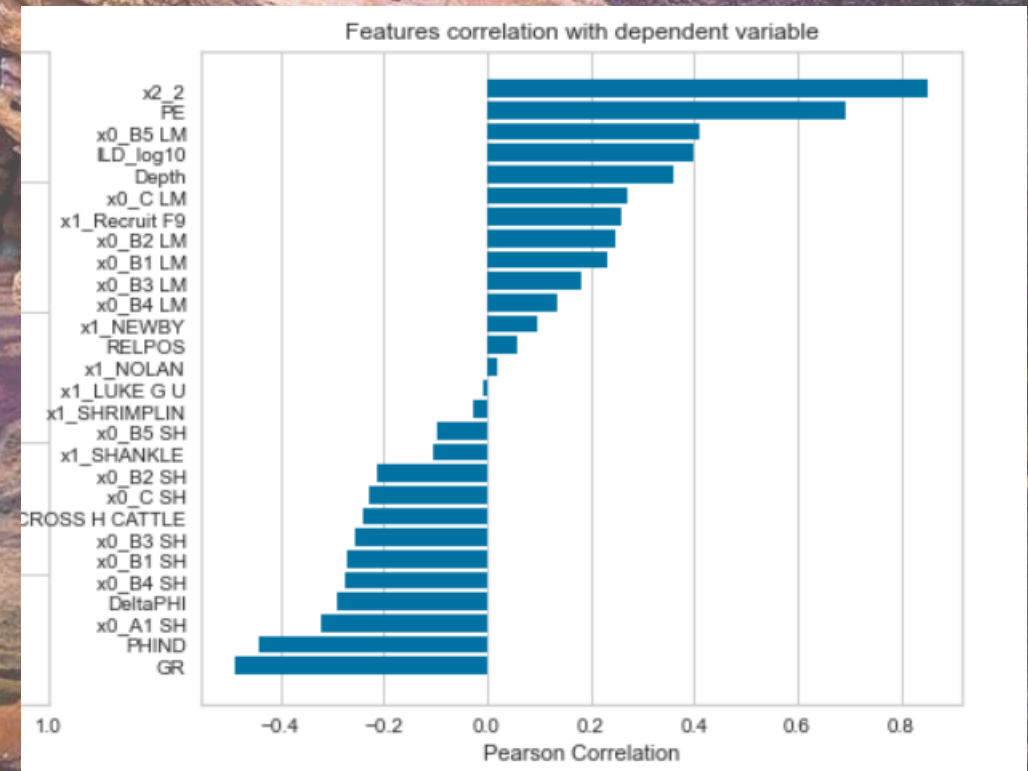
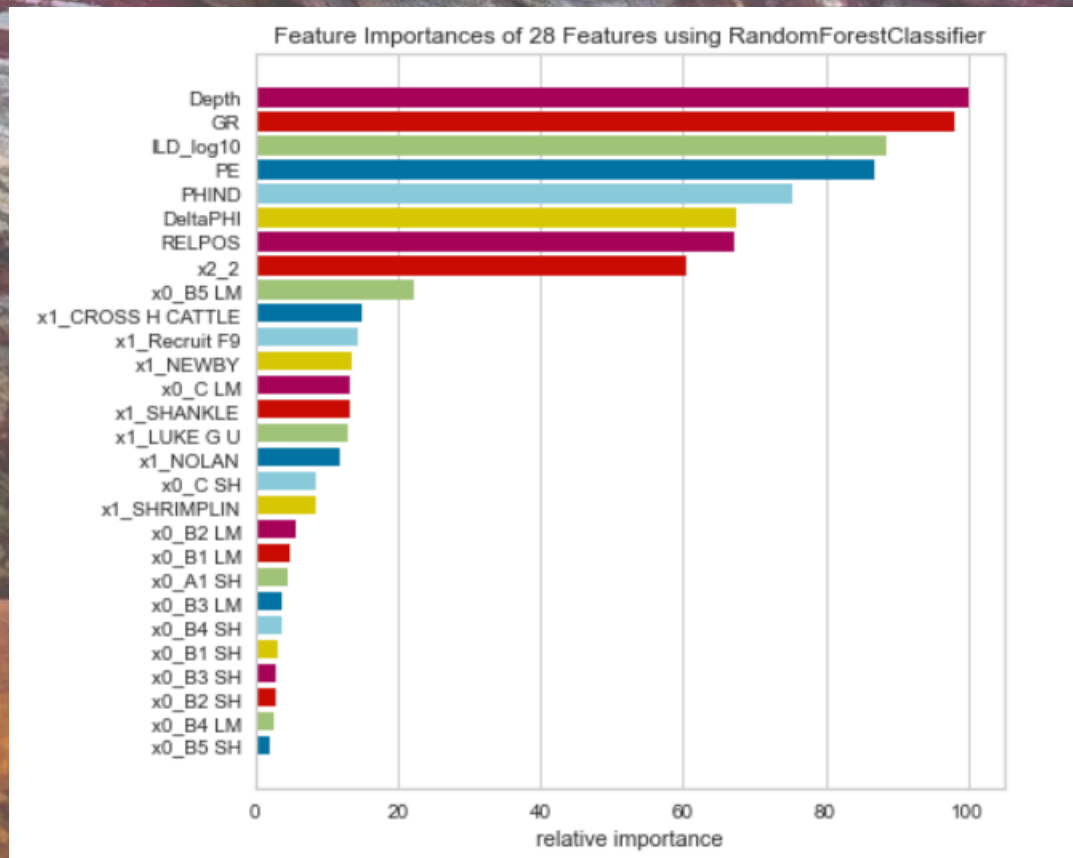
F1 - SCORE: 0.86



# Feature Rank and Importances



Feature Importance  
Feature Ranking

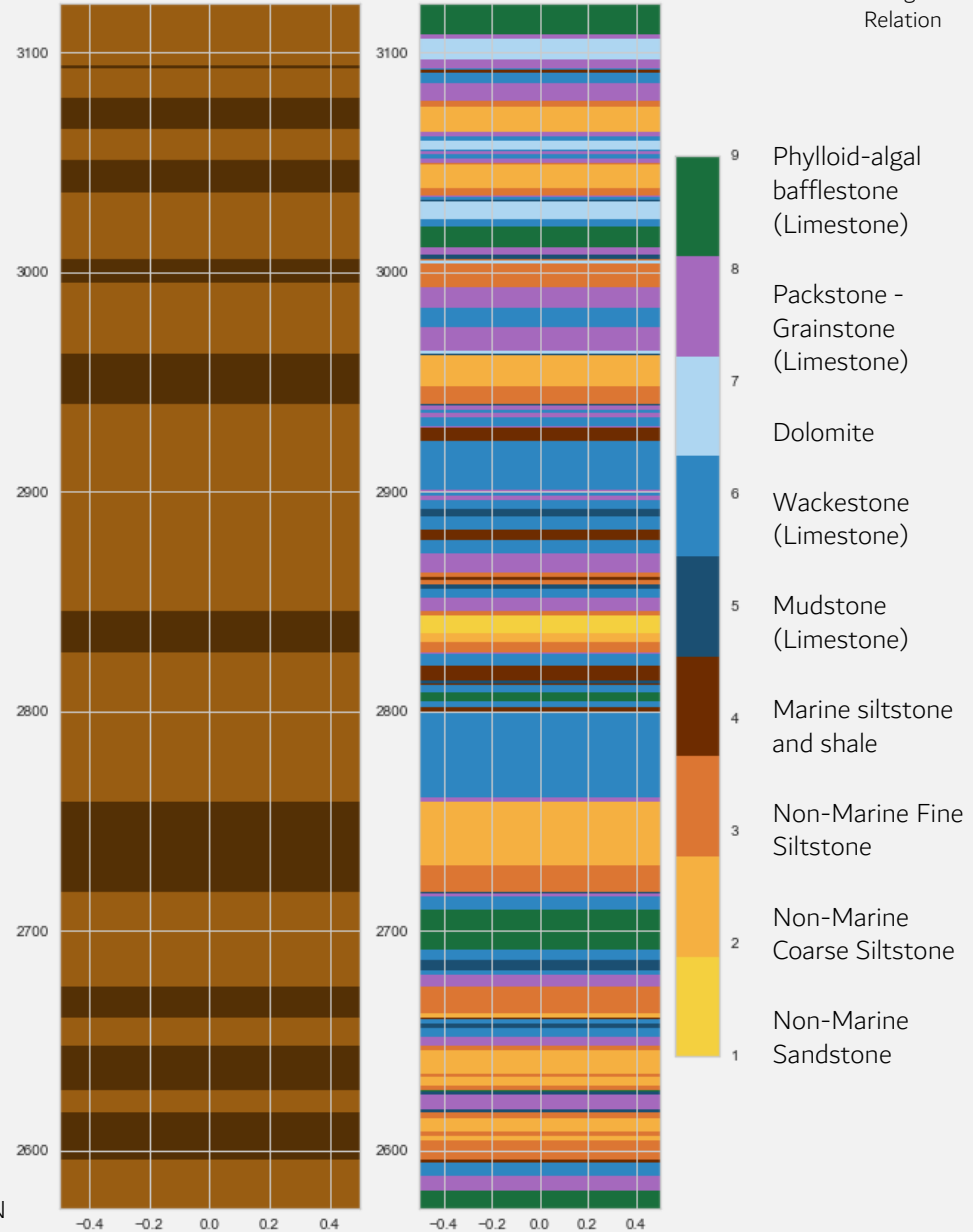


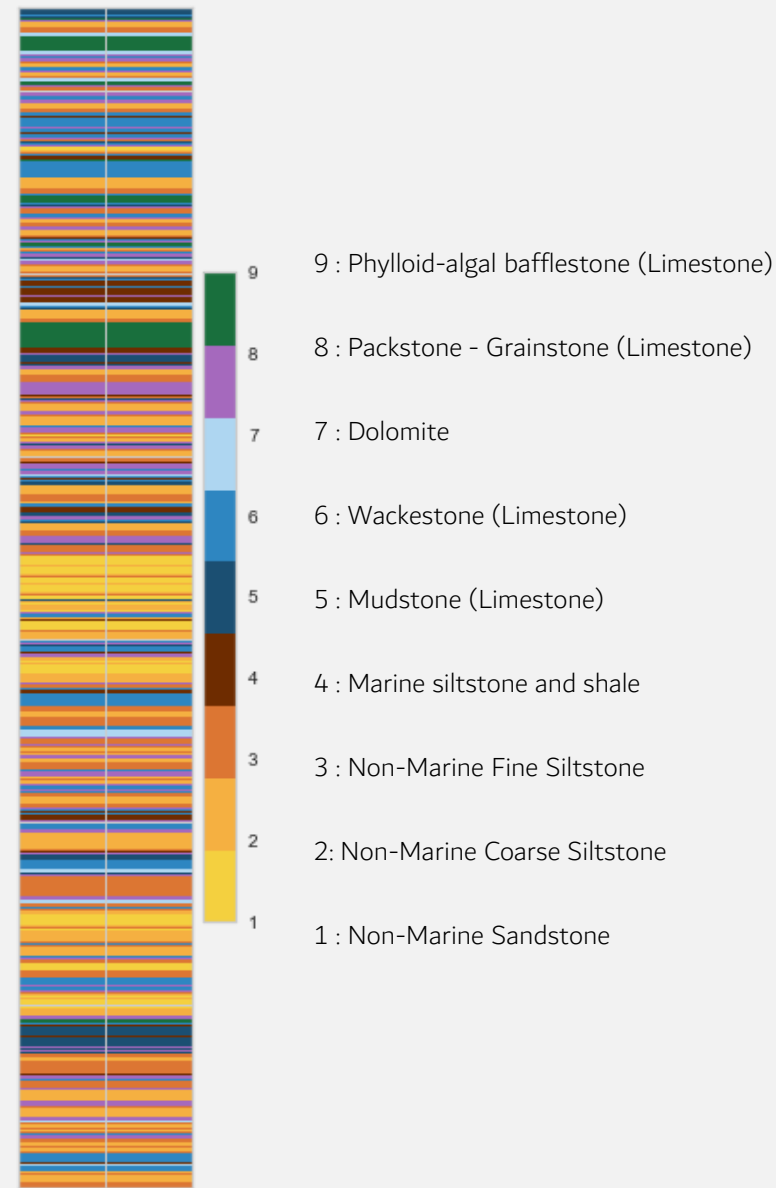
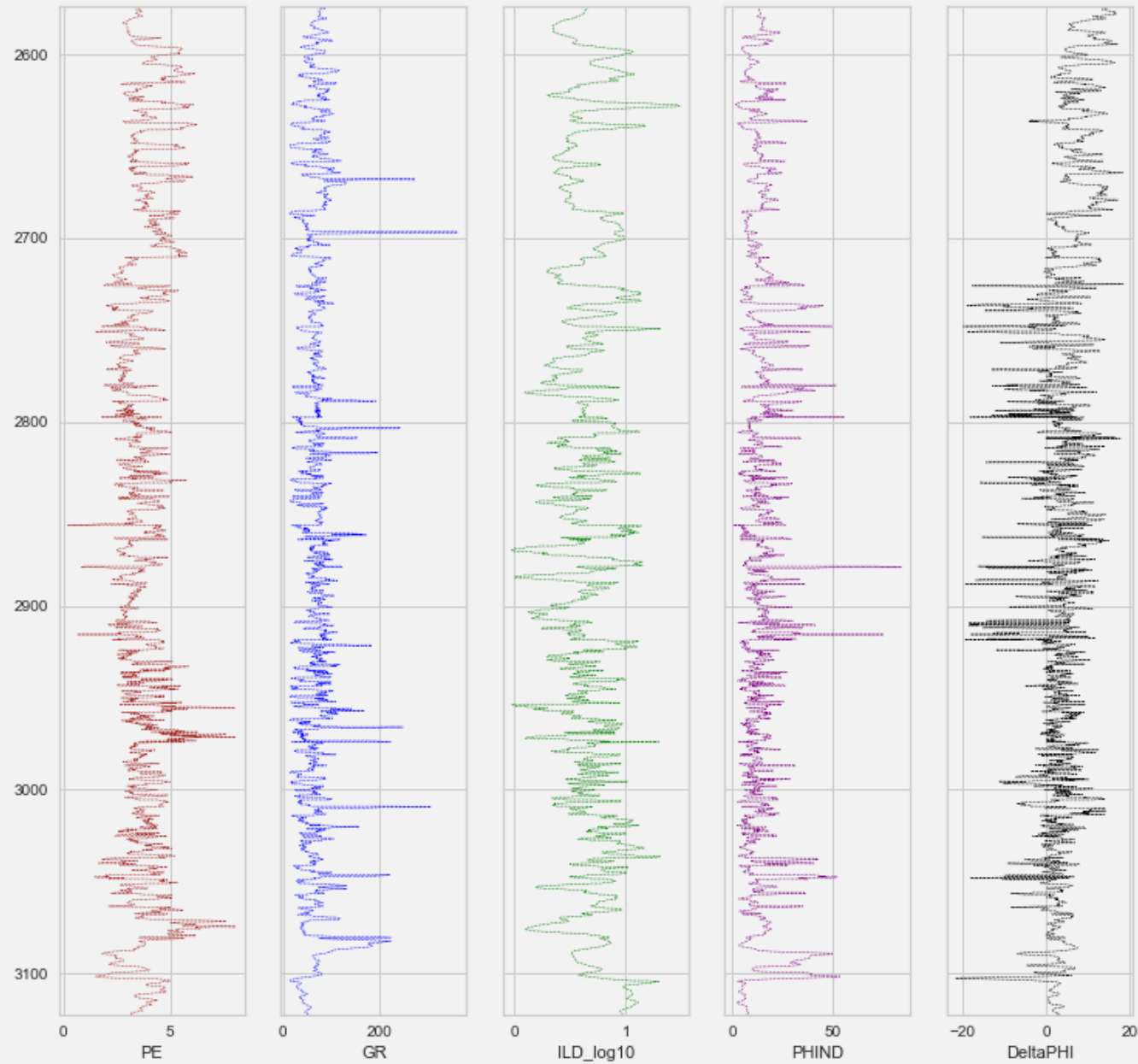


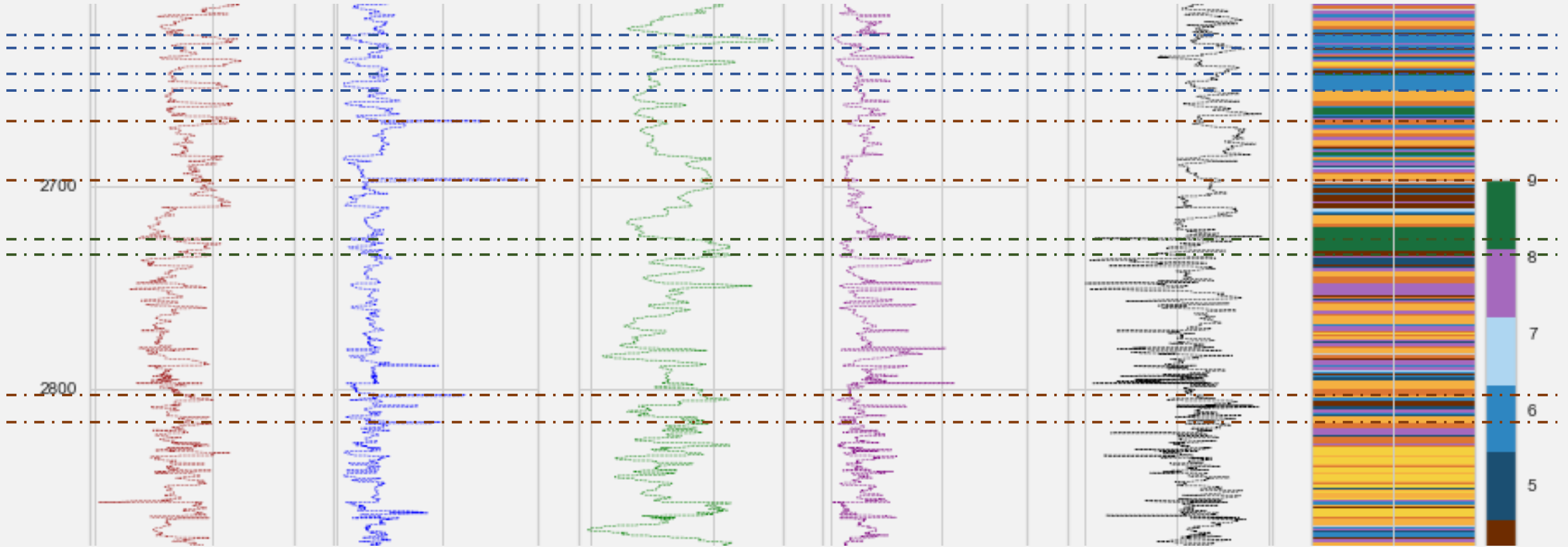
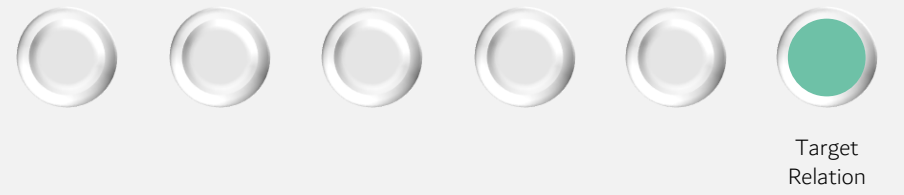
# Relation with Target



Marine  
Non-Marine





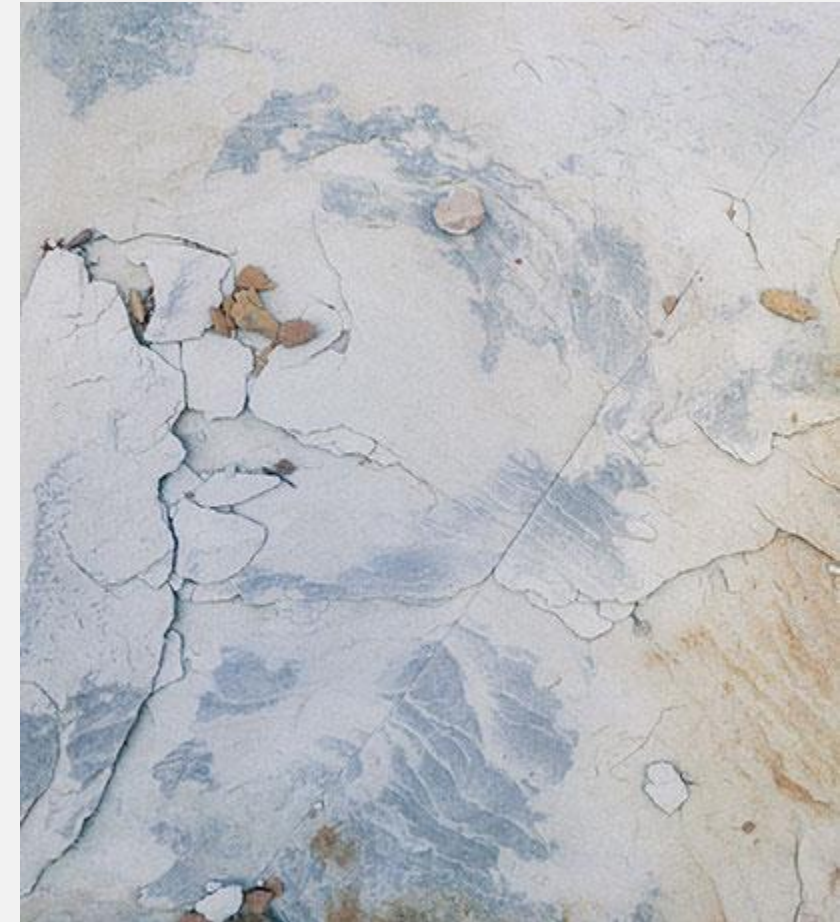


EXPANDING MACHINE LEARNING TO FORMATION EVALUATION

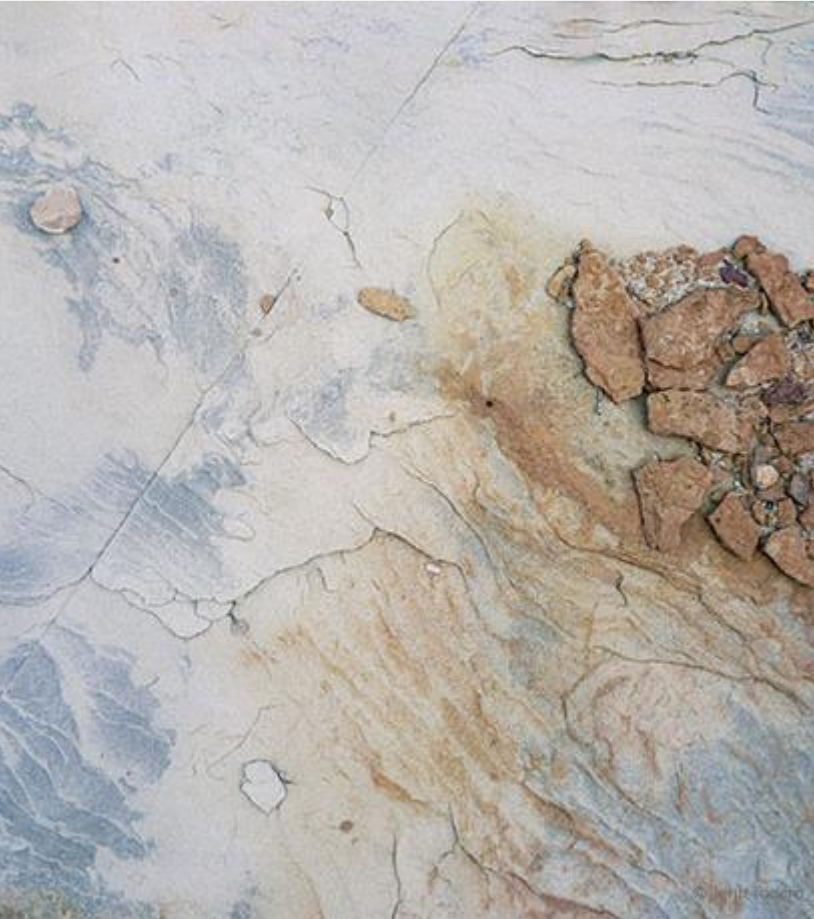


# Conclusions and Recommendations

- The best model took 11 SECS to classify the data into separate facies with F - 1 SCORE of 0.86.
- Machine learning model can be a very EFFICIENT and SPEEDY tool for facies classification compared to the cumbersome manual techniques currently used which take days to generate the results.
- The most important features that determine the accurate classification include the property of being MARINE or NON - MARINE, and the log values generated from PE, GR, N - D LOGS, RESISTIVITY and RELATIVE POSITION along with DEPTH. These show great influence since each of these values are unique to certain properties which define a facies.
- It is paramount, therefore, that these operations and data preprocessing is conducted meticulously before feeding in the data into the machine learning model.



# Future Work



- Further IMPROVING the model to include other methods of DISTANCE CALCULATION since distance is proved to be a major factor in the results.
- Analyze the EFFECT OF CLASS IMBALANCE to further improve our model.
- Expand and test this model for wells at DIFFERENT GEOLOGICAL LOCATIONS with other facies present to make this model applicable globally
- Use these PREDICTIONS AS A FEATURE in machine learning models to predict the main goal of facies classification.





# Thank You!

Questions?

