AI for inter-well saturation mapping

AI Berserkers

The problem and solution

Data Driven Method

Take the Data and use ensemble ML models:

-SciKit Bagging regressor. Train 2 hours. MSE - 0.086

- Catboost - CatBoost regressor.Train 10 minutes. MSE - 0.055

- XGBRegressor. Train 20 seconds. MSE-0.076

Data Driven+Physics

Take the Data and use ensemble ML models plus use physical models simulation:

- Archie's law, Archie's - Dahnov formula. We need more parameters for the equation.We use a=1,n=2,m=2 and Pickett plot.

Reformulated for electrical resistivity, the equation reads

$$R_t = a\phi^{-m} S_w^{-n} R_w$$

with R_t for the fluid saturated rock resistivity, and R_w for the brine resistivity.

Challenges deep-dive

Challenge 1

Data interpretation problem.

We try to understand how the data is organise. It takes some time to understand but we manage it.

Challenge 2

Materials about petrophysics.

We try to get some domain specific information about problem.

Challenge 3

Choose right models and combine with physics simulations.

We need to choose ensemble models, deep learning models. What best models Archie,double water ets?

Data Exploration

Porosity Distribution

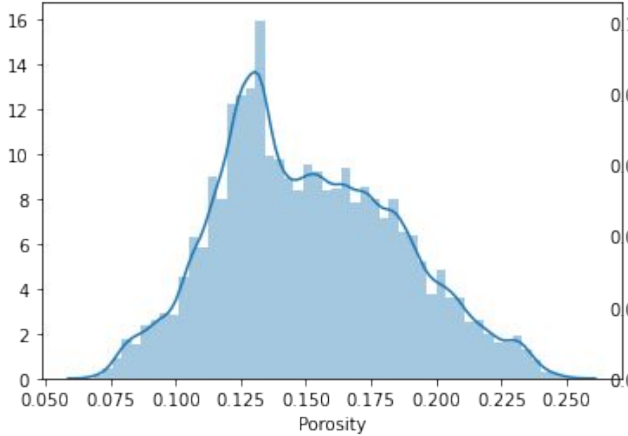
Clearly see close to normal data distribution.

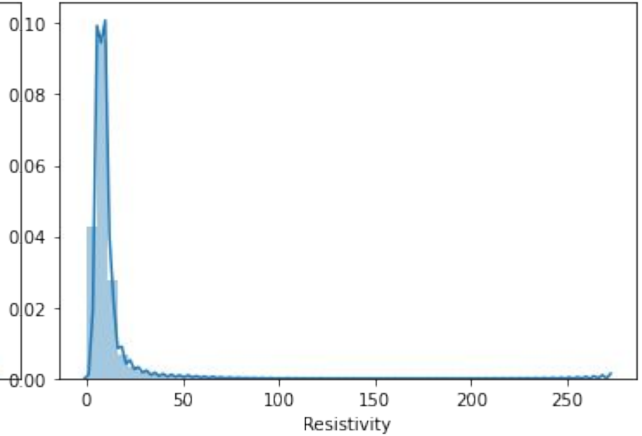
Resistivity Distribution

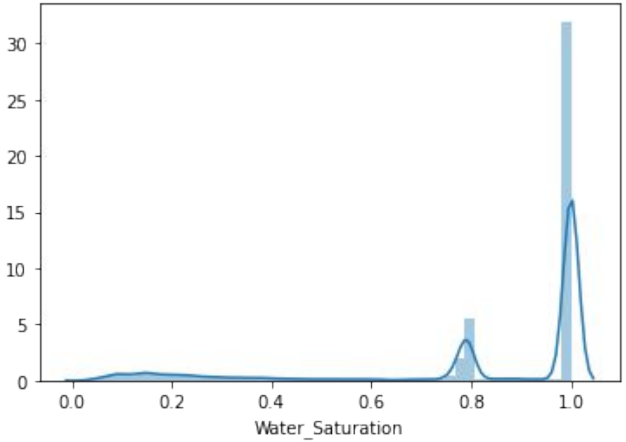
Most of the data between 0 and 15.

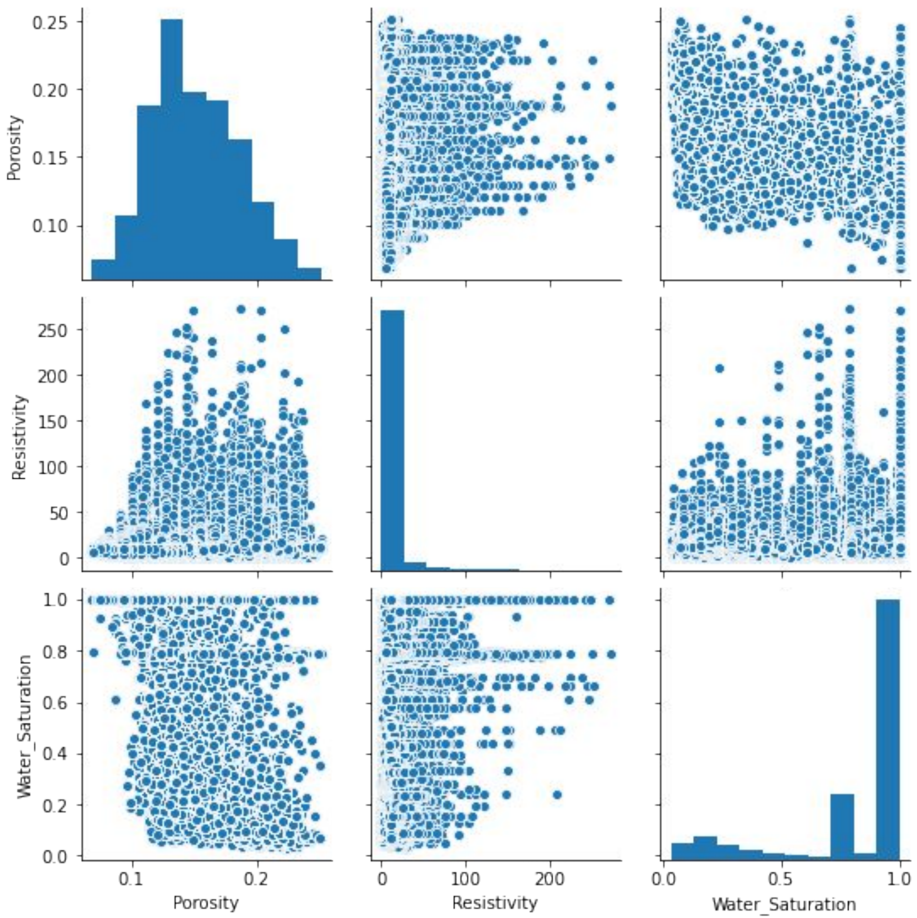
Water Saturation Distribution

Two peaks 1.0 and 0.8

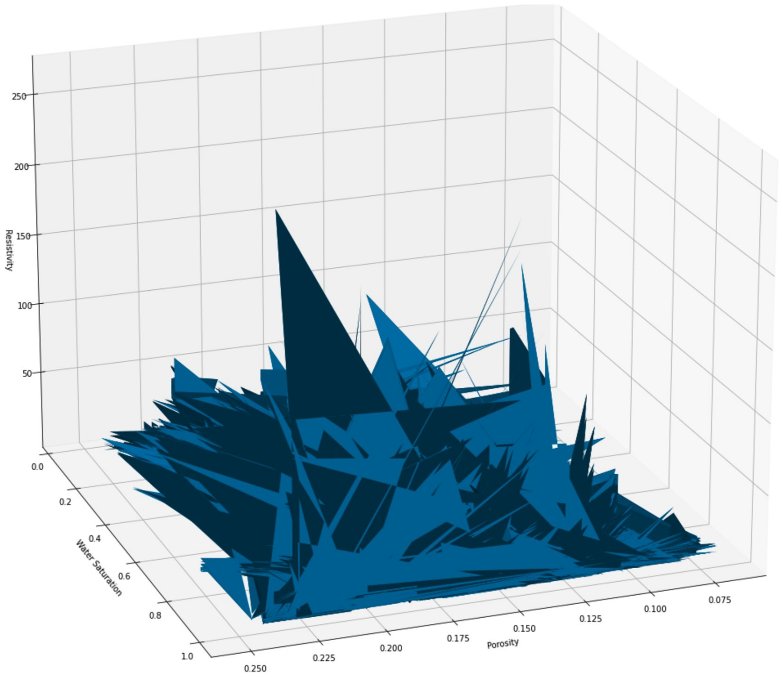






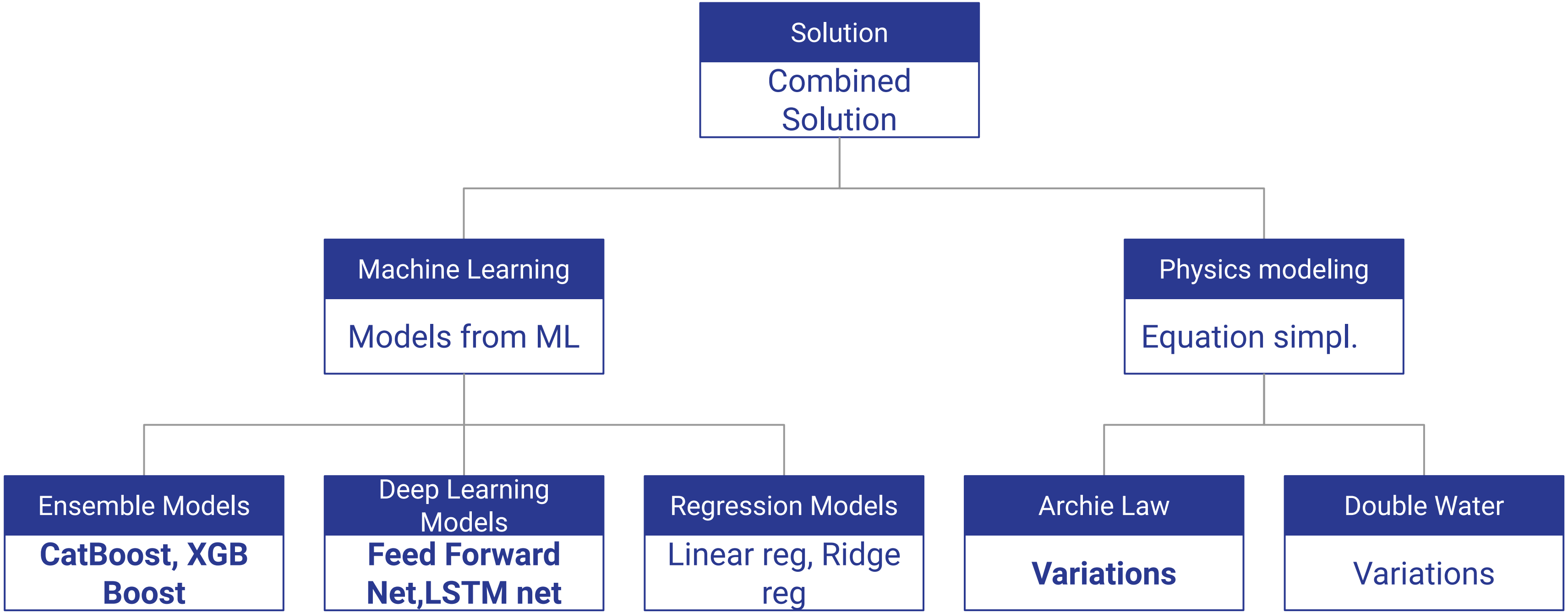


Pair plot correlation chart between all features. Not seen any good correlation.



3D Cube model.It helps us to correctly interpret the data.

Our Team Strategy



```
Model: "sequential_8"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
dense_24 (Dense)             (None, 256)               768
_________________________________________________________________
dense_25 (Dense)             (None, 128)               32896
_________________________________________________________________
dense_26 (Dense)             (None, 64)                8256
_________________________________________________________________
dense_27 (Dense)             (None, 1)                 65
=================================================================
Total params: 41,985
Trainable params: 41,985
Non-trainable params: 0
_________________________________________________________________
```

	loss	mae	mse	val_loss	val_mae	val_mse	epoch
5	0.074617	0.204733	0.074617	0.073983	0.195611	0.073983	5
6	0.074424	0.204024	0.074424	0.074939	0.193552	0.074939	6
7	0.074290	0.204298	0.074290	0.072857	0.205852	0.072857	7
8	0.074230	0.204226	0.074230	0.073058	0.209107	0.073058	8
9	0.074153	0.204061	0.074153	0.074732	0.187833	0.074732	9

Add Neural Net model.Same features - Porosity,Resistivity

	Porosity	Resistivity	Water_Saturation	Archie_model_sand	Archie_model_rand	Archie_model_rocks
count	180000.000000	180000.000000	180000.000000	180000.000000	180000.000000	180000.000000
mean	0.151212	10.228929	0.819488	0.544366	0.782127	0.512576
std	0.034774	10.673252	0.292216	0.189884	0.212872	0.198380
min	0.068508	0.050000	0.032581	0.064750	0.079026	0.047770
25%	0.125490	5.660000	0.787570	0.421163	0.636467	0.382580
50%	0.147530	8.330000	1.000000	0.518213	0.817362	0.481712
75%	0.176110	10.900000	1.000000	0.643062	1.000000	0.612276
max	0.251250	271.540000	1.000000	1.000000	1.000000	1.000000

We back to CatBoostRegressor because of the lower error.

```
pred_cat = cr.predict(X_test)
mean_squared_error(y_test, pred_cat)
```

```
0.054321294506122414
```

1: $SW = [(a / \varphi^m)(R_w / R_t)]^{(1/n)}$ classic general Archie equation for clean rocks (V/V)

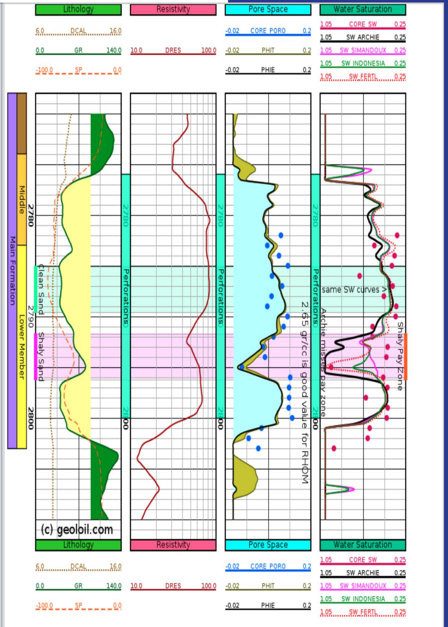
φ =	28	is a constant	is a curve number.	porosity $\varphi = \varphi_e = \varphi_t$ in clean rock (V/V)
R_t =	6	is a constant	is a curve number.	true uninvaded formation resistivity Ω.m
R_w =	25	is a constant	is a curve number.	water formation resistivity Ω.m
a =	0.81	is a constant	is a curve number.	tortuosity. Usually sands 0.81, carbonates 1
m =	1.7	is a constant	is a curve number.	sands around 1.8, carbonates around 2.0
n =	1.8	is a constant	is a curve number.	saturation exponent

2: SW by Fertl method for shaly sands: $SW = f(\varphi_e, VSH, R_t, R_w, a, m, \alpha)$

φ_e =	0	is a constant	is a curve number.	effective porosity (V/V)
R_t =	0	is a constant	is a curve number.	true deep, univaded resistivity (Ω.m)
R_w =	0	is a constant	is a curve number.	water formation resistivity (Ω.m)
VSH =	0	is a constant	is a curve number.	volume of shale (V/V)
a =	0.81	is a constant	is a curve number.	tortuosity. Usually sands 0.81, carbonates 1
m =	2.0	is a constant	is a curve number.	sands around 1.8, carbonates around 2.0
α =	0.30	is a constant	is a curve number.	empirical constant, usually 0.25 to 0.35

3: SW by Simandoux method for shaly sands: $SW = f(\varphi_e, VSH, R_t, R_w, R_{shale}, a, m)$

φ_e =	0	is a constant	is a curve number.	effective porosity (V/V)
R_t =	0	is a constant	is a curve number.	true deep, univaded resistivity (Ω.m)
R_w =	0	is a constant	is a curve number.	water formation resistivity (Ω.m)
VSH =	0	is a constant	is a curve number.	volume of shale (V/V)
R_{shale} =	4	is a constant	is a curve number.	resistivity of a nearby pure shale
a =	0.81	is a constant	is a curve number.	tortuosity. Usually sands 0.81, carbonates 1
m =	2.0	is a constant	is a curve number.	sands around 1.8, carbonates around 2.0



Notice that all the electrical equations shown above to estimate SW require to know the value for the formation water resistivity Rw. This is usually the most important parameter to estimate SW. The table below shows the most popular techniques aimed to estimate either salinity or its companion Rw value at the reservoir temperature:

	NaCl or Rw Technique	***Reliability***	***Source***	***Remarks***
1	*Ionic Water Analysis*	Best	Water sample	Water sample must be representative. Independent, log free technique
2	*Rw from water bodies and pockets*	Good	Logs	Needs to find 100% water bodies to work, like aquifers or water pockets
3	*Hingle Plot*	Good	Logs	Same math as Rw from H2O body, but from a cross-plot.
4	*Core SW vs. log SW match*	Medium	Logs and Core	Move salinity until match. Needs accurate lab SW measurements.
5	*Pickett Plot*	Poor	Logs	Rw & m estimation. Does not work if Phi is almost constant in water body
6	*SP Spontaneous Potential*	Worst	Logs	Last resource to try. Seldom yields accurate or usable Rw estimates

	↑ **Parameter increases**	***SW***	***SO or SG***
1	*a*	↑	↓
2	*m*	↑	↓
3	*n*	↑	↓
4	*Rw*	↑	↓
5	Salinity (*Rw* ↓)	↓	↑
6	*Rt*	↓	↑
7	φ	↓	↑
8	*Vsh* (φ_e ↓)	↑	↓
9	*Vsh* (Sat. correction)	↓	↑
10	*Rsh*	↓	↑
11	Fertl *a*	↓	↑
12	Grain density (φ ↑)	↓	↑
13	CEC *Qv*	↓	↑

	Parameter decreases ↓	***SW***	***SO or SG***
1	*a*	↓	↑
2	*m*	↓	↑
3	*n*	↓	↑
4	*Rw*	↓	↑
5	Salinity (*Rw* ↑)	↑	↓
6	*Rt*	↑	↓
7	φ	↑	↓
8	*Vsh* (φ_e ↑)	↓	↑
9	*Vsh* (Sat. correction)	↑	↓
10	*Rsh*	↑	↓
11	Fertl *a*	↑	↓
12	Grain density (φ ↓)	↑	↓
13	CEC *Qv*	↑	↓

We probably need better Rw then constant - 0.1, and it could be good to have Vsh.We need that for the new features Fertl and Semandu equation.