A Deep Learning Approach for Screening Enhanced Oil Recovery Methods

Jacob Pakula, Faculty Mentor: Dimitrios G. Hatzignatiou, Ph.D.
Department of Petroleum Engineering
UNIVERSITY of HOUSTON

Motivation/Background

What is Enhanced Oil Recovery?

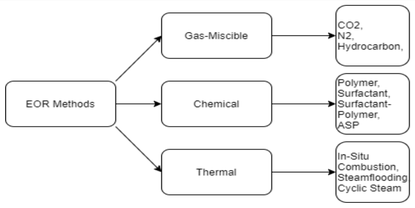
- Enhanced Oil Recovery (EOR) follows the secondary production stage of a hydrocarbon field and involves changing the physical properties of the reservoir fluid through an injected fluid
  - Goal: Increase Oil Displacement Efficiency
    $$E = E_D E_V$$
- There are 3 main EOR types
  - Chemical: Reduce IFT, Improve Mobility Ratio
  - Thermal: Reduce reservoir fluid viscosity
  - Gas-Miscible: Miscibility between both fluids

What is Deep Learning?

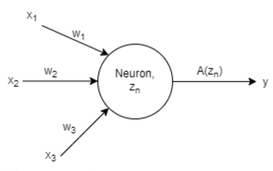
- Deep Learning is a class of Machine Learning algorithms that is based on the structure and functions of Artificial Neural Networks (ANNs)

What is the Goal/Motivation to Incorporate Deep Learning to EOR Projects?

- Rise of AI + Need to Lower Costs for E&P Companies + Future Energy Demands = Applying AI to EOR Projects
  - Reduce risks and costs associated with EOR Projects using Deep Learning
- Goal: Lay the foundations for researching and developing a model to screen for all EOR methods based of reservoir and fluid data



Techniques and Approach

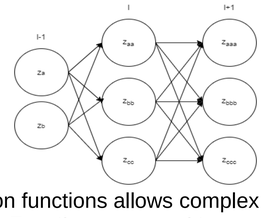


A single neuron: Receives a set of inputs (x) as well as a set of parameters in the form of weights (w) and biases (b) and computes a predicted output value which is passed through an activation function (A(z))

$$z = \sum_{i=1}^{n} (w_i \cdot x_i)$$

A layer of neurons: We use the following convention to compute the output of layer l in a network for each neuron:

$$z_i^{[l]} = w_i^T \cdot a^{[l-1]} + b_i$$
$$a_i^{[l]} = A^{[l]}\left(z_i^{[l]}\right)$$



Activation and Loss Functions: Activation functions allows complex relationships from data to be learned. Loss Functions are used to show the progress of the learning process. Since we have multiple classes of EOR methods represented in the dataset, I used the categorical cross-entropy loss function.

$$CE = -\sum_{i}^{C} t_i \log(A(z)_i)$$

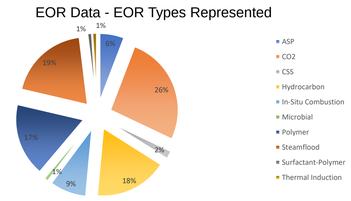
Backpropagation: Find the loss function's minimum → For each iteration: Calculate the values of the loss function partial derivatives w.r.t. each of the parameters of our neural network

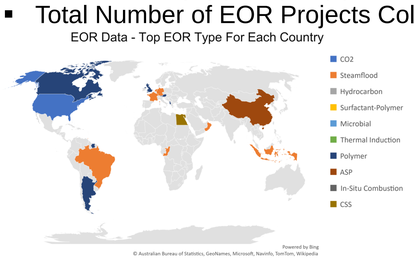
$$dW^{[l]} = \frac{\partial L}{\partial W^{[l]}} = \frac{1}{m} dZ^{[l]} A^{[l-1]T} \qquad dA^{[l-1]} = \frac{\partial L}{\partial A^{[l-1]}} = W^{[l]T} dZ^{[l]}$$
$$db^{[l]} = \frac{\partial L}{\partial b^{[l]}} = \frac{1}{m} \sum_{i=1}^{m} dZ^{[l](i)} \qquad dZ^{[l]} = dA^{[l]} \cdot A'(Z^{[l]})$$

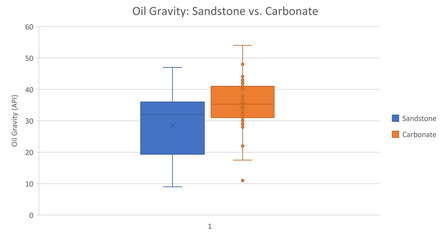
Results and Analysis

EOR Data Gathering and Analysis:

- Collected EOR Data
  - Oil Properties → Gravity (°API), Viscosity (cp), Composition
  - Reservoir Characteristics → Field Type, Oil Saturation (%PV), Porosity (%), Permeability (md), Temperature (°F), Pressure (psi), Net Pay (ft), Depth (ft)
- Total Number of EOR Projects Collected → 103



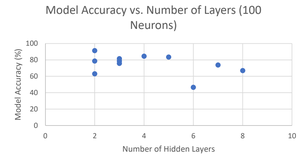




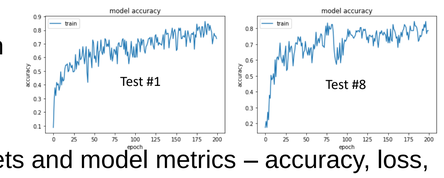
- Made Boxplots comparing the oil properties and reservoir characteristics between two types of hydrocarbon-bearing reservoirs: Sandstones and Carbonates

3 Questions to Investigate:

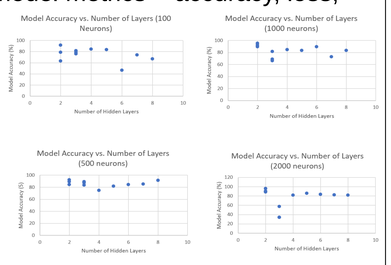
1. Relationship between number of hidden layers and model metrics – accuracy, loss, predictions
   - Model Accuracy → A negative trend in accuracy for cases with more hidden layers
   - Test #1 – 2 Hidden Layers with 100 Neurons Total
     - Logarithmic trend with mid-fluctuations in accuracy
   - Test #8 – 5 Hidden Layers with 100 Neurons Total
     - Large drop in accuracy around epochs 75-85

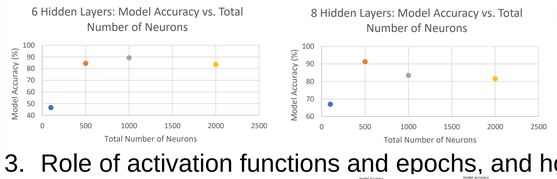


   - Both test cases were able to predict the correct test case EOR method within 90% accuracy
   - Analysis → Model accuracy reaches a maximum accuracy threshold range around 2-4 Hidden Layers in the neural network



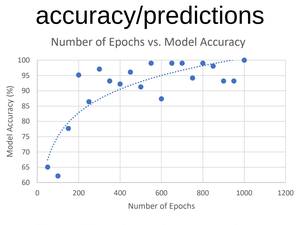
2. Relationship between total number of neurons/nets and model metrics – accuracy, loss, predictions
   - 100 Neurons
     - Avg Accuracy = 74.935
     - STD = +- 11.715
   - 500 Neurons
     - Avg Accuracy = 85.791
     - STD = +- 4.742
   - 1000 Neurons
     - Avg Accuracy = 82.437
     - STD = +- 9.065
   - 2000 Neurons
     - Avg Accuracy = 74.051
     - STD = +- 21.000

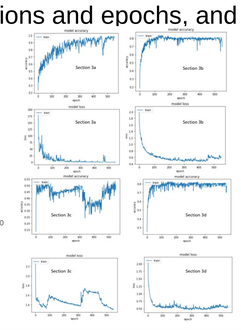




- Left: 6 out of the 8 curves with different number of hidden layers shows the model accuracy reaches a peak value around 500-1000 total neurons

3. Role of activation functions and epochs, and how it affects training and model accuracy/predictions



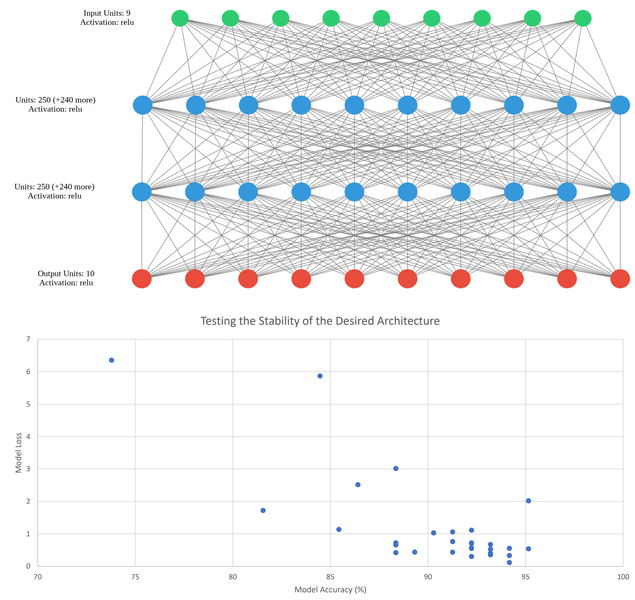


- Top: Testing the number of epochs from 50-1000
- Section 3a: 1st HL = 2nd HL = Relu Activation Function
  - Avg Accuracy and STD = 96.312, +-2.706
  - Avg Loss and STD = 0.227, +-0.332
- Section 3b: 1st HL = 2nd HL = Sigmoid Activation Function
  - Avg Accuracy and STD = 79.418, +-3.601
  - Avg Loss and STD = 0.562, +-0.095
- Section 3c: 1st HL = Relu, 2nd HL = Sigmoid
  - Avg Accuracy and STD = 50.873, +-3.766
  - Avg Loss and STD = 1.269, +-0.110
- Section 3d: 1st HL = Sigmoid, 2nd HL = Relu
  - Avg Accuracy and STD = 76.796, +-1.669
  - Avg Loss and STD = 0.573, +-0.061

Conclusions

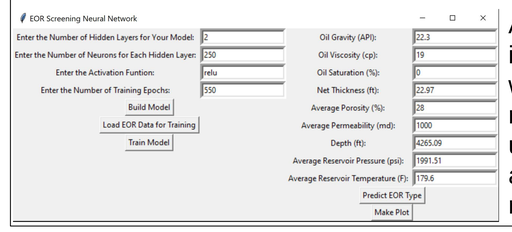
Neural Network Model Architecture:

- 2 to 3 Hidden Layers
  - Model accuracy decreases with an ever-greater number of hidden layers
- 500-1000 Total Neurons in the Neural Network
  - Model loss begins to decrease with a greater total number of neurons
    - Dilution of EOR data
- Both Hidden Layers Utilize the Relu Activation Function
  - Greater model accuracy than the other activation functions
- 500 Epochs for Training
  - Model accuracy seems to reach its peak after 500 epochs
    - More epochs might cause over-training of the EOR dataset



EOR Screening Model – User Interface

I've built a graphical user interface that allows a user to create their own EOR neural network and then train and make predictions with their model on what EOR method to use based on inputted reservoir and fluid data.



Allows a brief intro into DL without the need for understanding algorithms and math.

Acknowledgements

A special thanks to Dr. Dimitrios G. Hatzignatiou for his guidance this summer on the research project, and the Summer Undergraduate Research Fellowship (SURF) program for providing me the funds to work on this research during the Summer 2021.