Machine Learning Application to A Physical System
Abdullah Alakeely

1. Problem Statement: Predicting the behavior of subsurface fields is a challenging task that require accurate knowledge of the physical properties of the field in question, and the development of complex numerical simulation models that describe the behavior of the system.

2. Task: Given rate profiles, and derivatives as a time series of T time steps, from a specific reservoir as an input feature, represented by \( [Q, Q' , \Delta Q'] \in \mathbb{R}^{T \times 3} \), with corresponding pressure profiles as an output, represented by \( [P, Q, P', Q'] \in \mathbb{R}^{T \times 4} \).

Input

Al Model

Output

3. Data: 9 different rate and pressure profiles are generated randomly using fully-implicit solution of a single reservoir under water injection, with only one producer. Every profile is considered one example. We start by one example for training, one for dev set, and the rest for testing.

4. Model Architecture: Number of nodes and learning rate is found through grid search.

Cost Function: MSE

Performance Metric: Average sum of error in prediction where:

\[ \text{Error} = |y - \hat{y}| \]

5. Initial Search: A grid search of \( n = \{0.0001, 0.001, 0.01, 0.1\} \), \( 11 \) epochs using 2 examples in training/2 in dev/ the rest for testing.

6. Switch Sequence: every 5 epochs using 2 examples in training/2 in dev/ the rest for testing.

7. Regularization and SGD: every 5 epochs using 2 examples in training/2 in dev/ the rest for testing.

8. Learning: SGD vs Adam

Adam

SGD

9. Examples of Prediction

10. Conclusions and Future Work:
- This work demonstrated the nature of iterative process to arrive at the right architecture and data. Grid search is helpful, but a random search should be more efficient and will be implemented as an improvement.
- Amount of data the network sees during training affect the performance. K-fold cross validation should be used to reach the best division of data that generate the best performance with matching data distributions.
- Optimizing using SGD with gradient clipping worked better than Adam in this case. This indicate the uniqueness of every problem and what works best in one case might not work in other. SGD will be explored further in the future.
- LSTM and GRU could be helpful and should be explored further.

References: