A Novel Methodology to Detect Frac Hits by Implementing Deep Learning and Distributed Data

VU NGUYEN

Objective: Develop deep learning algorithm to detect frac hits

Outlines:

- Introduction
- Data Acquisition/Data Wrangling
- Data Visualization
- Deep Learning Sparse Autoencoder
- Conclusions

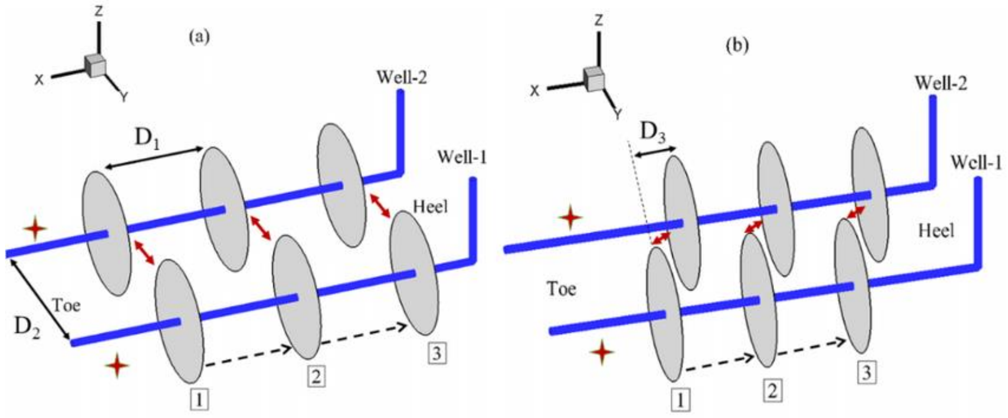
Introduction

- The advance in horizontal drilling and hydraulic fracture significantly increase the United States oil and gas production

- Operators tries to drill their horizontal wells as close as possible to maximize production.

- As a result, hydraulic fracture hits among the horizontal wellbores have become norm.

- Frac hits were reported to have positive and negative impacts.



Data Acquisition/Data Wrangling

Data Wrangling

- There are 79 hydraulic fracture stages
- 68,333 LAS files. Each LAS file is 1 second measurement
- Each LAS file was loaded into Pandas DataFrames
- Filtered out 0-8,000 ft depth
- Separated the negative and positive slow strain
- Filtered data into pre-frac (5 minutes before the stage starts), during frac and

post frac (after the stage ended to the 5 minutes before the next stage start)

Data Acquisition/Data Wrangling (cont.)

Engineer Features

- 16 features were computed
- Cumulative positive/negative strain pre-frac
- Cumulative positive/negative strain during-frac
- Cumulative positive/negative strain post-frac
- Total strain during-frac
- Total strain after-frac
- Difference in positive strain between during-frac vs. post-frac
- Difference in negative strain between during-frac vs. post-frac
- Root-mean-square (RMS) was applied to cumulative positive/negative strain pre-frac, during-frac

post-frac

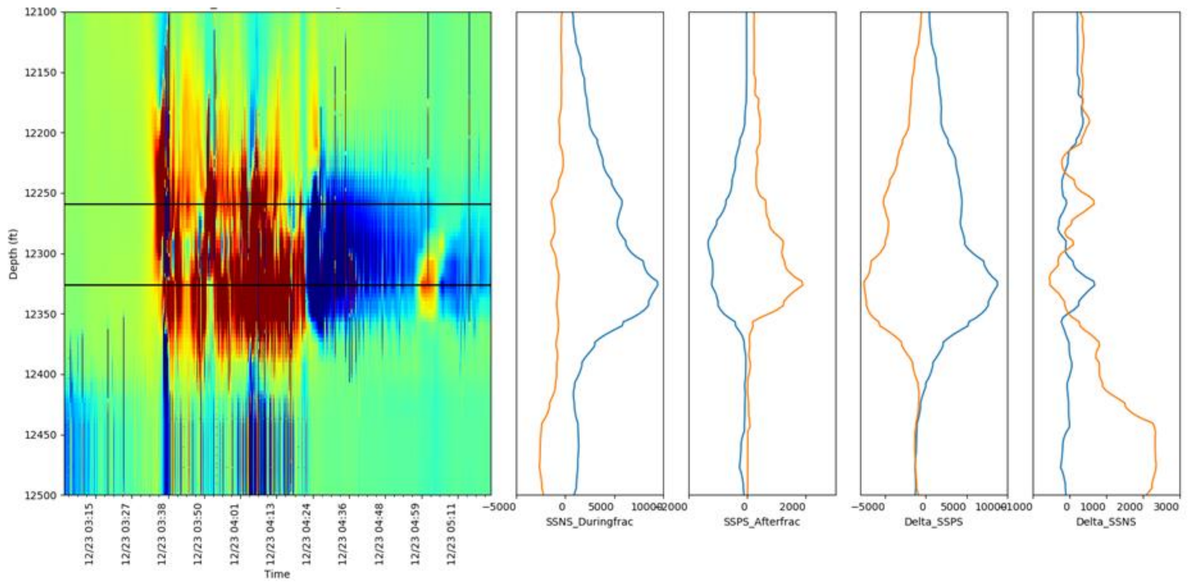
Data Acquisition/Data Wrangling (cont.)

<u>Data Labels</u>

- Initially, frac hits labels were picked by engineers in the field

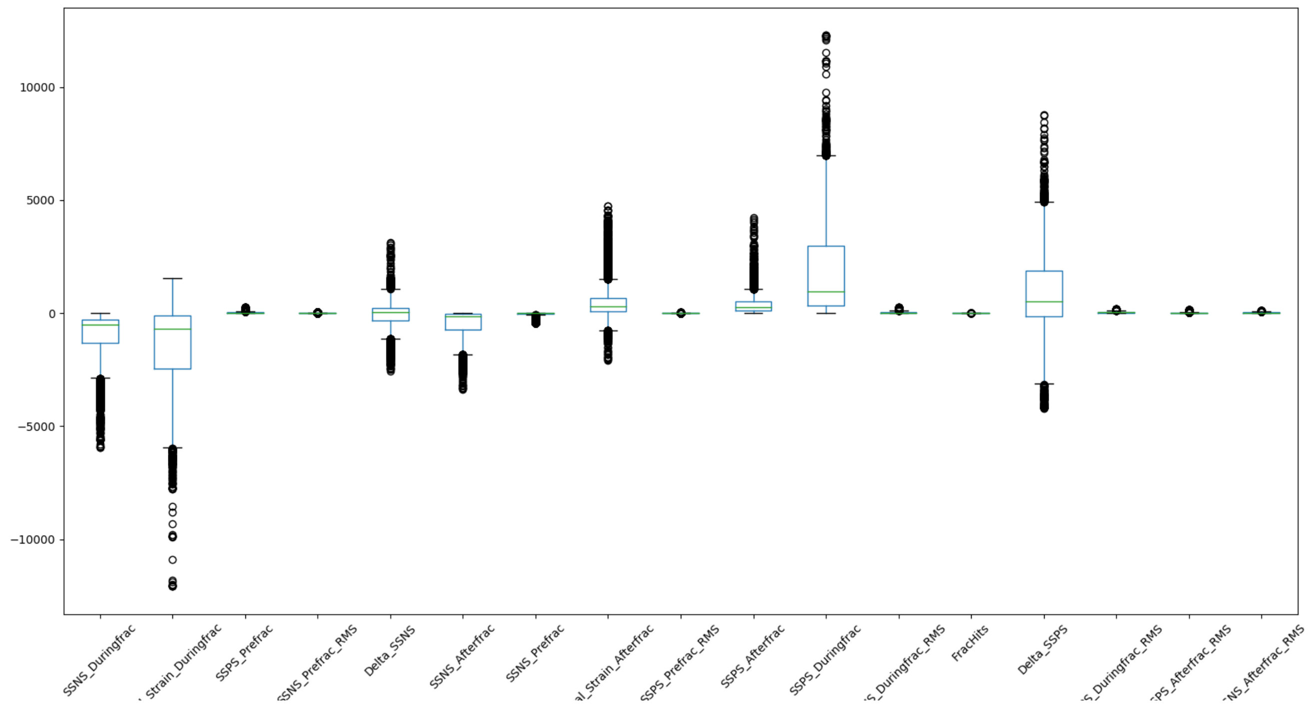
- Highly inconsistent and subjective

- Developed a methodology to pick the frac hits based on the correlation between peaks



Data Visualization

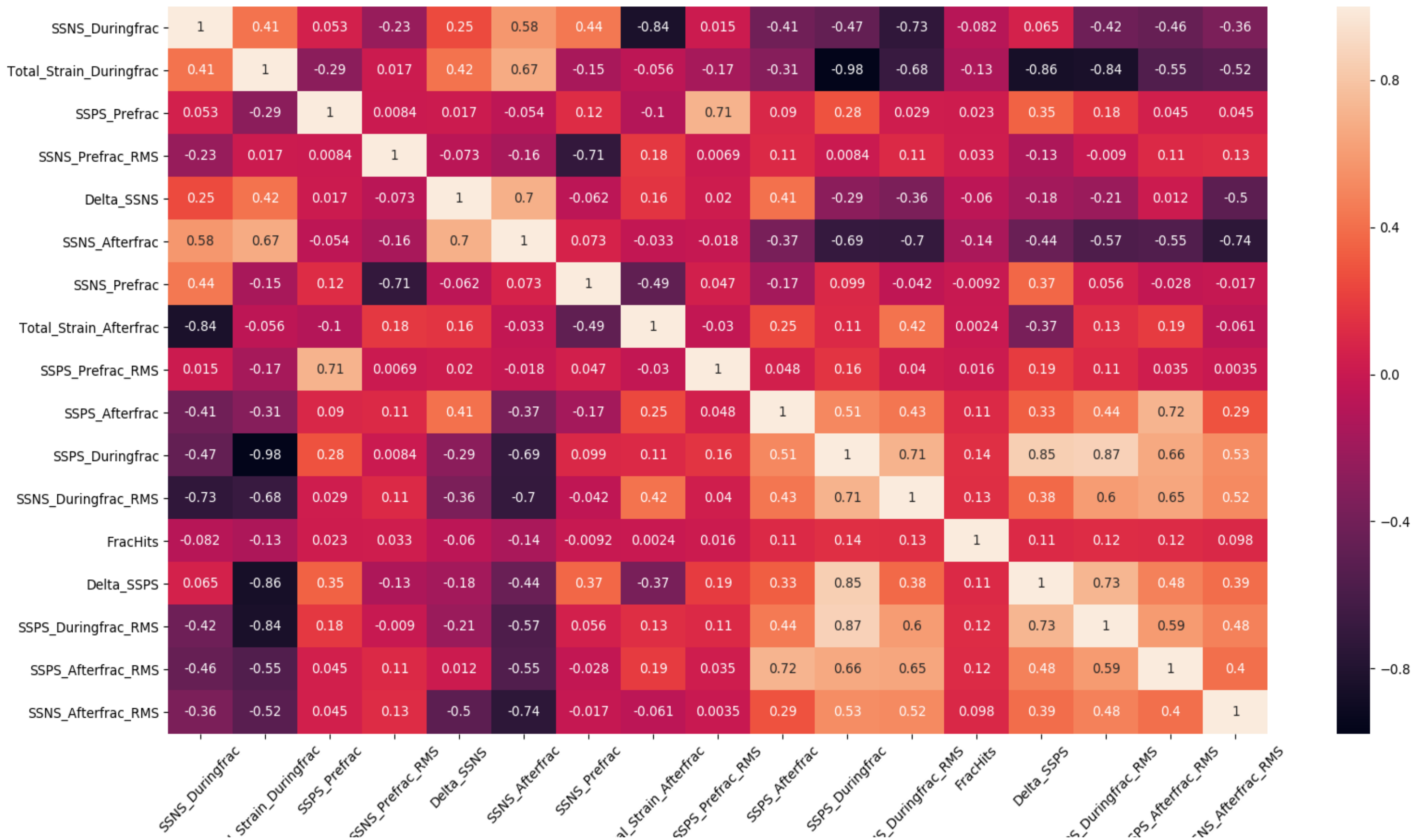
- All the feature data have outliers
- Leave outliers as they are because they may reflect the true nature of measurement data



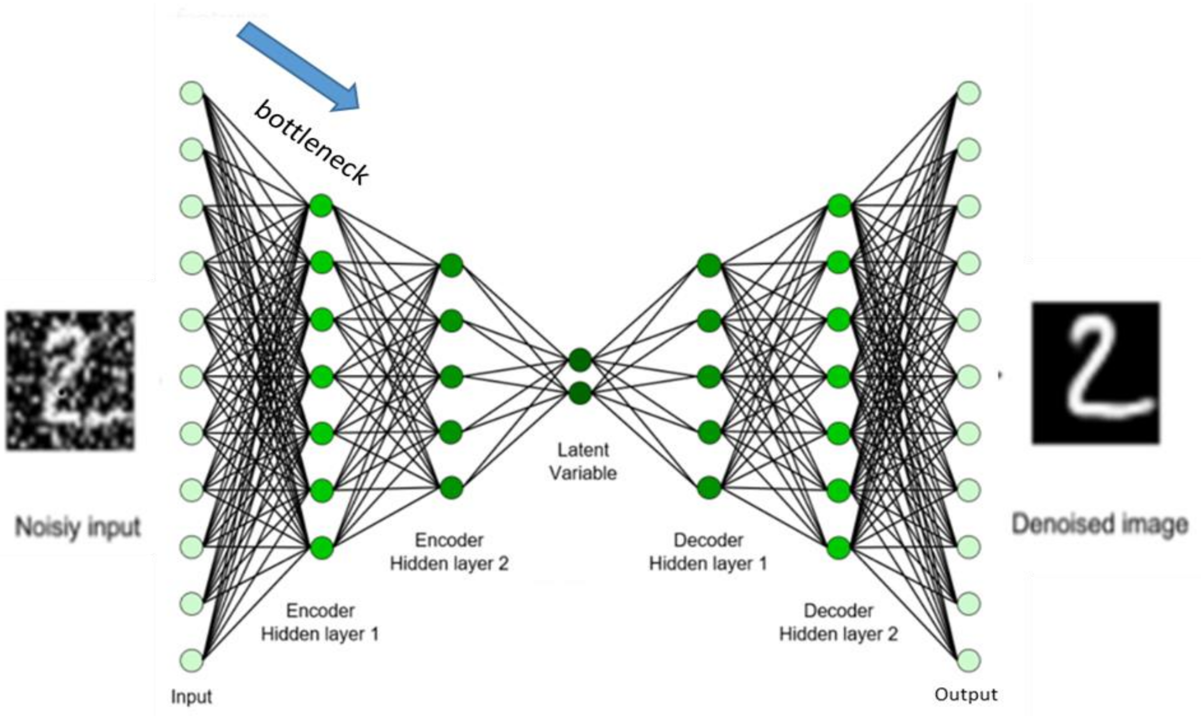
Data Visualization (cont.)

- Cumulative negative strain after-frac exhibits negative linear relationship with cumulative positive strain during-frac

- Total strain during-frac shows strong correlation with cumulative negative strain after-frac and inverse correlation with cumulative positive strain during-frac and Delta-SSPS



Sparse Autoencoder - Background



Source: Medium

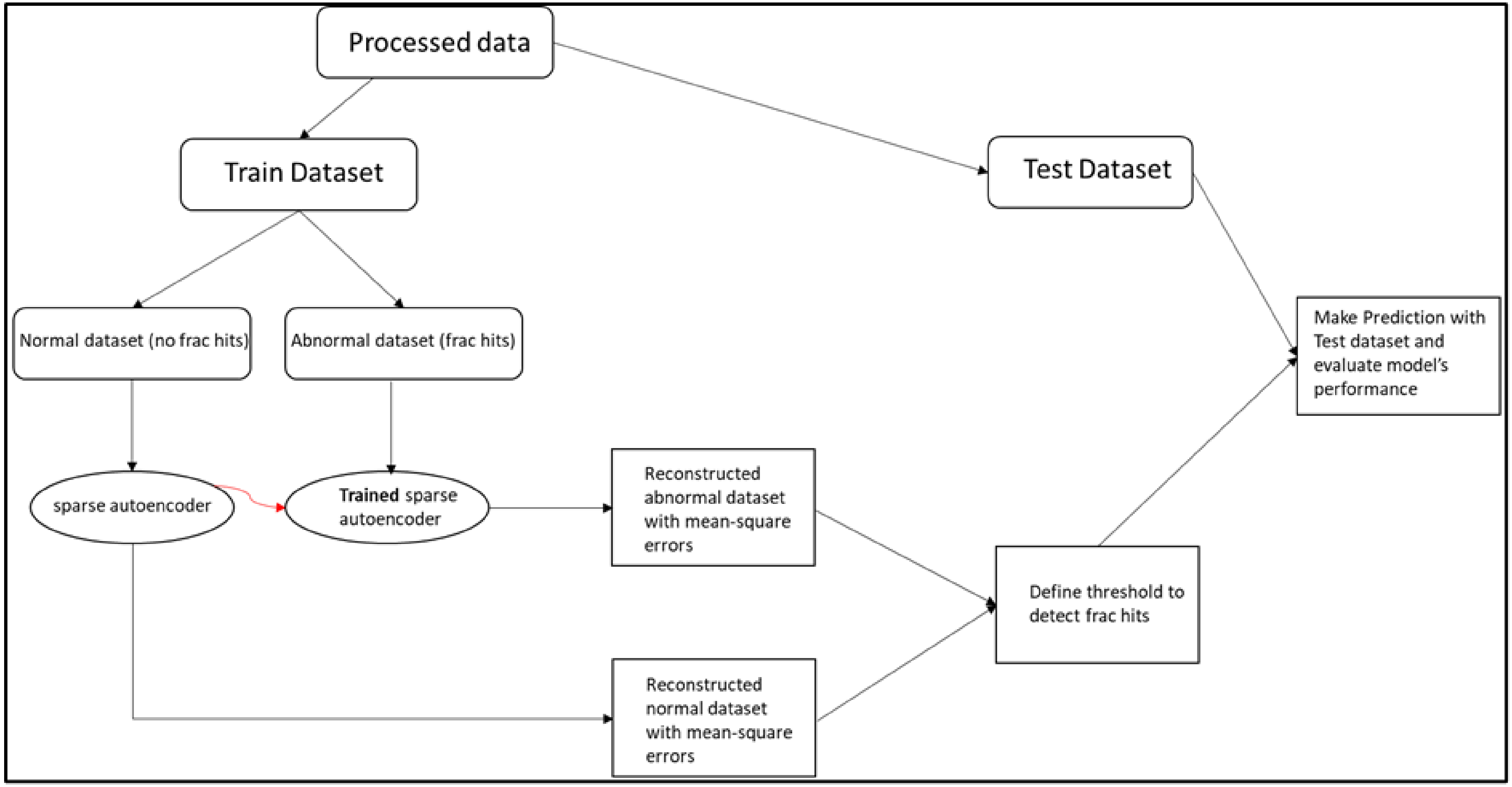
- **Autoencoder** is a special type of deep learning, which takes high dimensional input vector and pass it through to hidden layer(s) called **encoder** to compress the high dimensional input vector to lower dimensional vector. The lower dimensional vector is called **latent vector**.

- The latent vector is then goes through another hidden layer(s) called decoder to create the output vectors, that is very close to the original input vector without the noise.

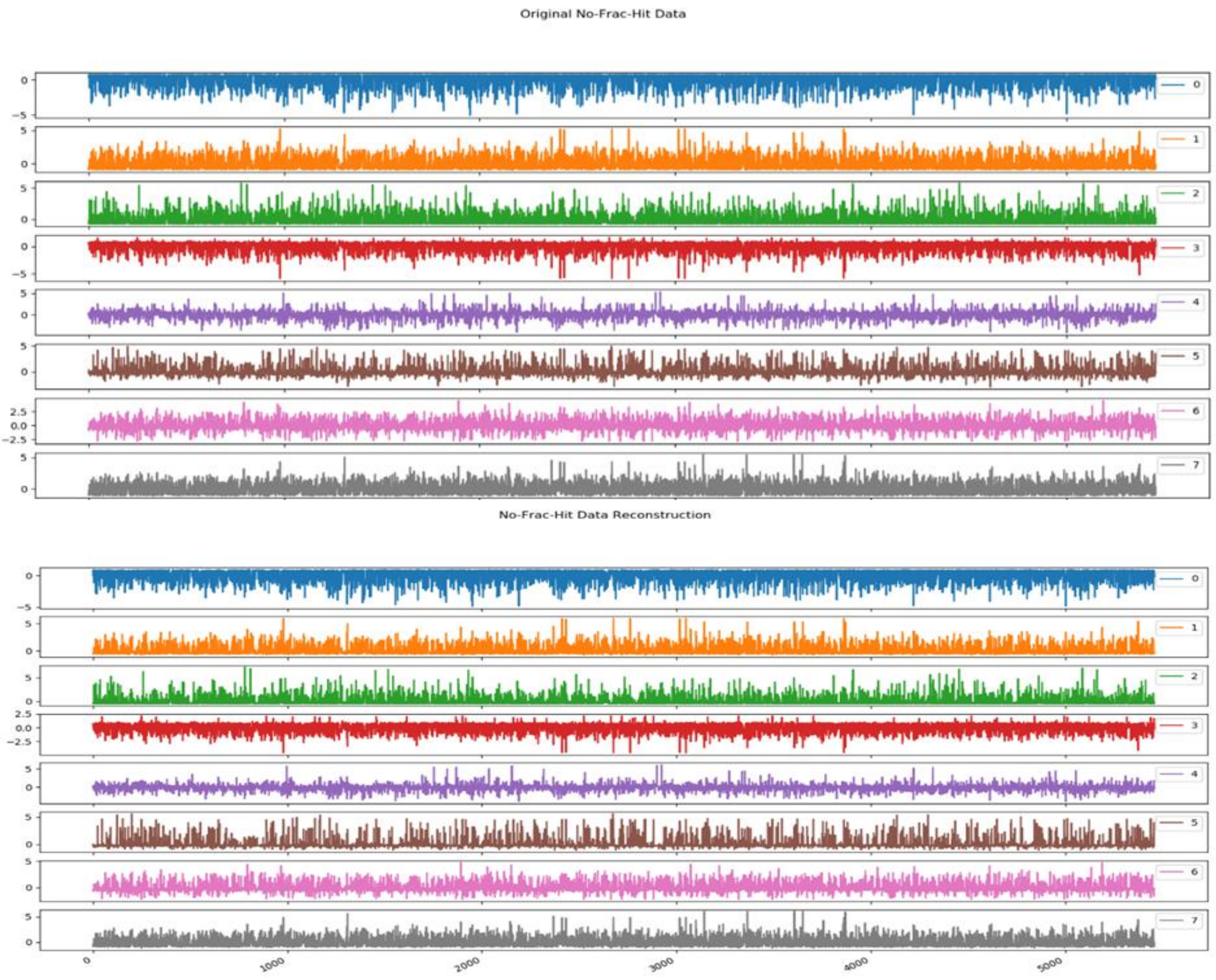
- Note that, the number of neurons in encoder layer(s) is less than input vector's features to create the bottleneck in the network, so the network is forced to learn the main features of the input vector, not the noise. As a result, the output images from the left showed clearer digit images than the input images.

- In this study, **sparse autoencoder** is used to detect frac hits. Sparse autoencoder is simply an autoencoder in which the bottleneck is created by penalizing activation of hidden layers instead of having fewer neurons than input vector's features.

Sparse Autoencoder - Methodology

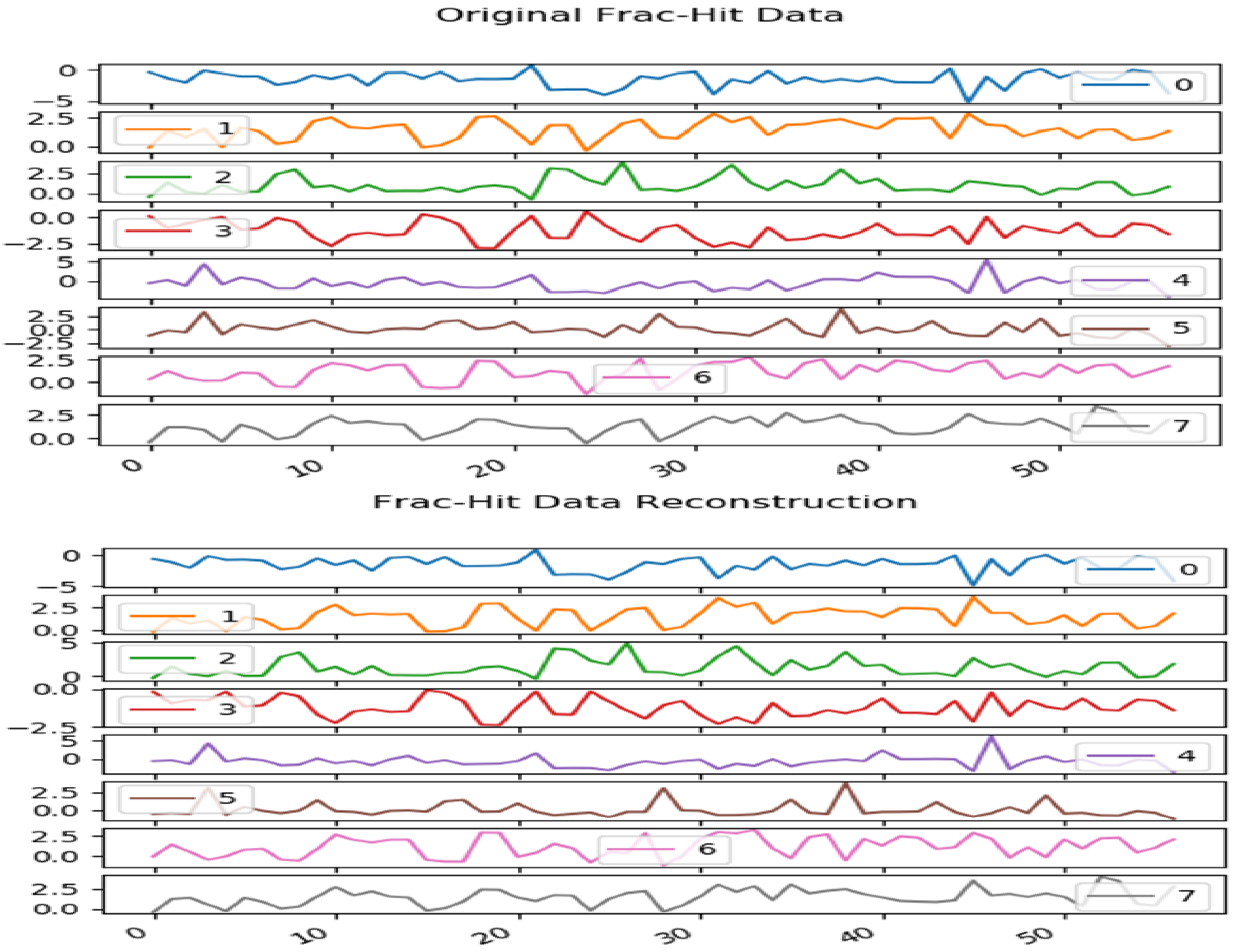


Sparse Autoencoder – Normal Input Data Reconstruction



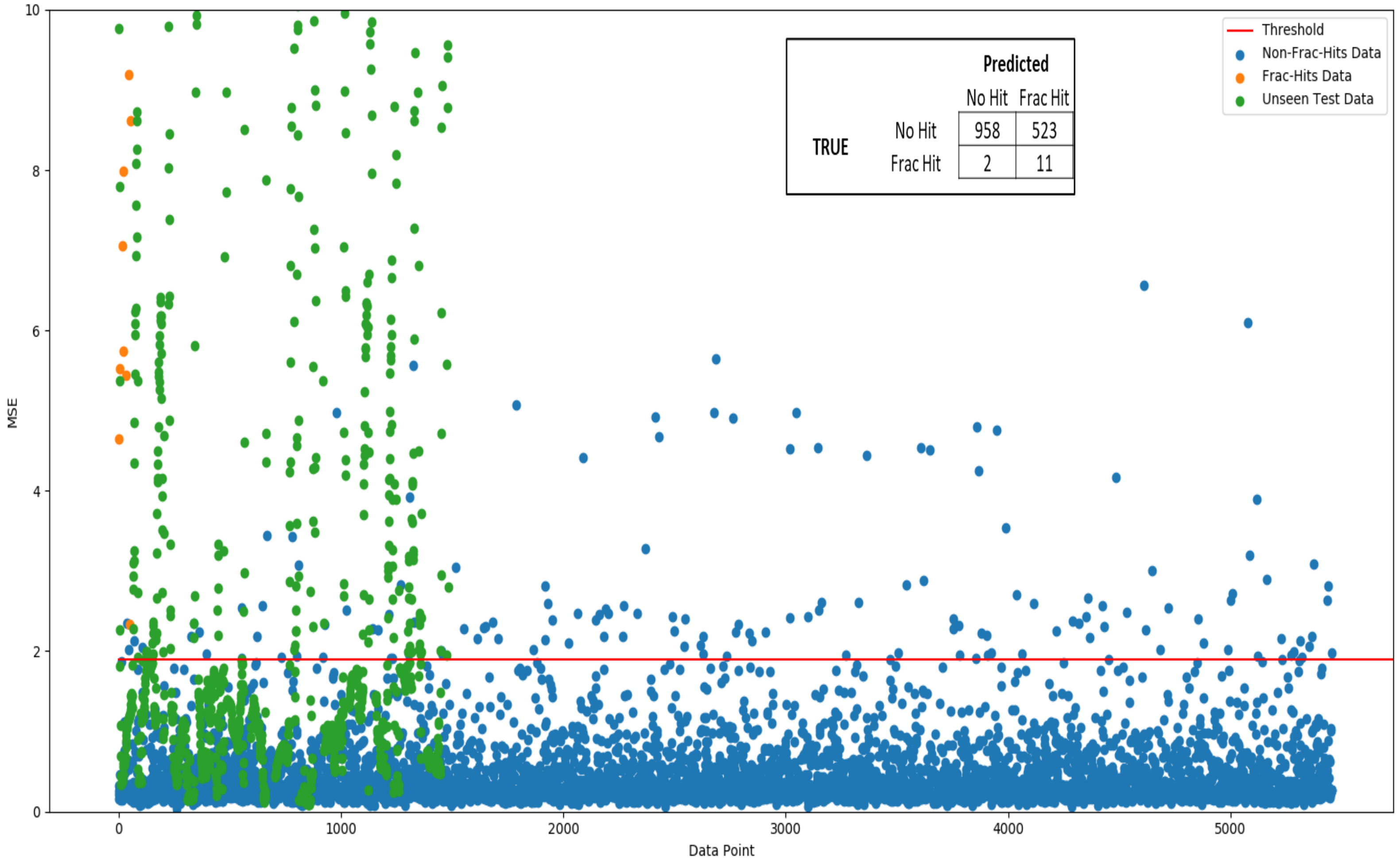
- As shown, the No-Frac-Hit Data Reconstruction resembles many features of original data
- Model is efficient to learn the original input data.
- Therefore, the MSE between original no-frac-hit data and its reconstruction should be small.

Sparse Autoencoder – Normal Input Data Reconstruction



- The Frac-Hit Data Reconstruction shows similar features as original data. However, the magnitudes are different.
- Therefore, the MSE between original frac-hit data and its reconstruction should be higher.

Sparse Autoencoder – Results



		Predicted	
		No Hit	Frac Hit
TRUE	No Hit	958	523
	Frac Hit	2	11

- A threshold was defined. Any data point that has MSE higher than the threshold will be considered as frac hits.
- Any data points below the threshold will be considered as non-frac-hit.
- Model can correctly detect 11 frac hits out of 13, which yield 85% accuracy.
- It only can detect 958 non-frac hits out of 1481, which yield 65% accuracy.
- There is a trade-off of accuracy between frac hit and non-frac hit events.

Conclusions

- Sparse autoencoder was developed to detect the frac hit events. With the defined threshold, the model can predict the frac hit events with 85% accuracy.
- However, there is a tradeoff between the false alarm vs. the accuracy to detect frac hits. While the model is very efficient to detect the frac hit events, but it can only predict the non-frac hit events with 65% accuracy.
- Investigate further to understand why several non-frac hit events exhibits very high MSE.

Future Work

- Pursue to develop LSTM autoencoder and other machine learning algorithms to deal with time series data. So, the time of frac hit events can be identified.