Abstract

The oil and gas industries play a crucial role in meeting human needs, and with the advancement of technologies, a sensor is now used to monitor the distribution of oil and gas. Any unwanted consequences can be avoided by analyzing the results shown by the sensor. Dealing with said problems can be challenging, making machine learning an invaluable tool for this task. This paper uses two deep learning approaches—Gated Recurrent Unit (GRU) and basic Recurrent Neural Network (RNN)—to construct autoencoder models for detecting anomalies in natural gas pipeline data. The dataset itself consists of 8590 data points that were gathered by sensors in a natural gas pipeline for 1 year that were made into hourly format. Both models will be trained using the said dataset to aim for minimal reconstruction errors. We compare their performance across five different architectural configurations using mean squared error (MSE) to identify the most effective setup. After getting the optimal model, we compare the original and reconstructed data to calculate the errors using Euclidean distance and set the anomaly threshold accordingly based on that. By determining the threshold value, we can detect anomalies in the data. Qualitative analysis reveals that both models perform well. The GRU method gives a slightly better result than RNN. The only slight difference may be due to the complexity and size of the dataset. Further studies of these methods using varying data volumes and complexity are warranted to understand the relative strengths of each model.

Keywords: anomaly detection, autoencoder RNN, GRU, gas pipeline