Abstract

The natural gas pipeline network plays a vital role in global energy distribution,This study discusses the performance comparison of two machine learning models, Extreme Gradient Boosting (XGBoost) and Long Short-Term Memory (LSTM), in predicting pressure in natural gas pipeline transmission networks. The background highlights the importance of pressure monitoring for ensuring the safety and efficiency of energy distribution. The dataset consists of 61,315 pressure measurements collected from August 1, 2020, to July 31, 2021, with the variables used including Normalized Pressure, IC5, Volume Rate, Energy Rate, and Temperature. In the methodology, exploratory data analysis (EDA) was carried out to understand the characteristics of the data, followed by preprocessing steps that included normalization and handling of missing values. The XGBoost model demonstrated excellent performance, reducing the Mean Squared Error (MSE) from 0.000392 to 0.000312 after hyperparameter tuning and improving the coefficient of determination (R^2) from 0.986478 to 0.989233. The LSTM model also showed improvement, with the MSE decreasing from 0.0009221 to 0.0009003 and R^2 increasing from 0.98366 to 0.98405. The evaluation results indicate that XGBoost provided more accurate predictions compared to LSTM, although LSTM excelled in capturing more complex temporal patterns. This study emphasizes the importance of selecting the appropriate model based on data characteristics and operational requirements. Suggestions for future research include exploring other models, developing more advanced features, and applying hybrid approaches to improve prediction accuracy. The integration of the prediction system with a real-time visualization dashboard is also recommended to support better operational decision-making.

Keywords: Gas Pipeline Network, LSTM, Machine Learning, Pressure Prediction, XGBoost