Abstak

Gas pipeline networks are essential for the safe and efficient distribution of gas to various locations, but they are also vulnerable to numerous technical issues, with gas leaks being one of the most dangerous. Gas leaks in pipelines can lead to catastrophic outcomes, including fires, explosions, and significant environmental harm. Early detection of these leaks is therefore crucial to prevent such severe consequences. This research focuses on developing a robust anomaly detection method for gas pipeline networks using an ensemble-based machine learning approach, specifically through random forest and gradient boosting algorithms. The study highlights the critical importance of early detection of gas leaks in pipeline infrastructure to prevent catastrophic consequences, including fires, explosions, and environmental damage. Leveraging extensive operational pipeline datasets from oil and gas companies, the research begins with a comprehensive data preprocessing phase designed to ensure the highest level of data quality and integrity. Both random forest and gradient boost models are rigorously implemented and trained on this dataset, with a focus on clustering data into decision trees or groups to effectively identify anomalies. The primary objective is to compare the accuracy of the random forest and gradient boost models while also exploring the potential for enhanced performance by combining these two powerful methods. The effectiveness of the anomaly detection system is meticulously evaluated using F1-score and accuracy metrics, which provide a clear measure of model performance. This research aims to significantly improve the safety and reliability of gas distribution systems by delivering a cutting-edge machine learning approach for anomaly detection in gas pipelines. The study's results, demonstrating an accuracy of 0.90 and an F1-score of 0.90, indicate strong and reliable performance