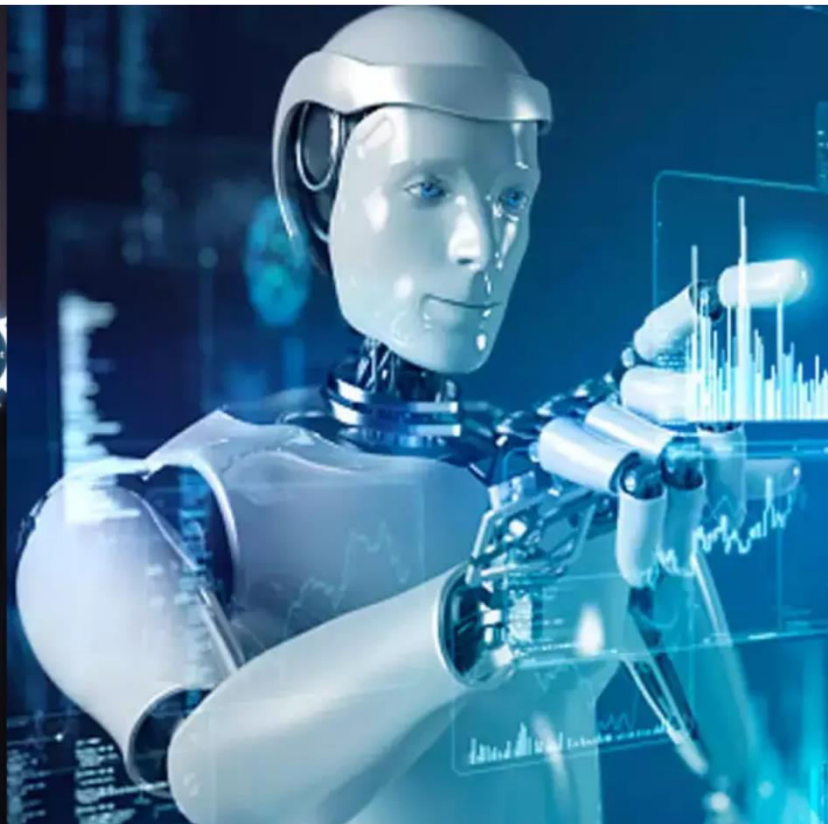




International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.699

Volume 14, Issue 5, May 2025

Smart Anomaly Detection with Real-Time Analytics and Chatbot Support

Pratham Umesh Pai, Shashank S Sangalad, Sneha N S, Sathvik S Kerudi, Shilpa Mathpathi

Department of Computer Science and Information Technology, REVA University, Bangalore, India

ABSTRACT: The steel industry, one of the largest global energy consumers, faces challenges in managing energy consumption for efficiency, cost reduction, and sustainability. This paper presents a real-time anomaly detection system using machine learning to identify irregularities in energy usage specific to steel manufacturing. The system analyses data including energy usage (kW), currents, CO2 emissions, and timestamps for scalable detection.

The system uses the Isolation Forest algorithm to isolate anomalies in high-dimensional data, with real-time processing to flag deviations immediately. It features a user-friendly interface for anomaly visualization and an integrated chatbot for interaction with non-technical stakeholders, providing notifications and actionable recommendations. This research demonstrates the benefits of combining machine learning with interactive interfaces in industrial settings. The solution offers scalability, robustness, and usability, addressing challenges in energy monitoring and promoting sustainability. Future work will expand the system's capabilities to predictive analytics and enhanced chatbot features.

KEYWORDS: Energy Monitoring, Machine Learning, Isolation Forest, Real-Time Anomaly Detection, High-Dimensional Data, Chatbot Interface, Energy Consumption, Industrial Efficiency, Sustainability, Operational Data.

I. INTRODUCTION

Anomaly detection is crucial in energy-intensive sectors like steel manufacturing, where identifying irregularities in operational data can improve efficiency and reduce costs. This project focuses on designing a real-time anomaly detection system using machine learning to monitor energy consumption and operational data. Traditional methods are inadequate for handling high-dimensional, dynamic datasets. This research builds on existing work in the field. Kwon and Han's study [1] introduced anomaly detection using self-organizing maps, demonstrating neural networks' potential in large datasets. This inspired the selection of machine learning techniques for the detection framework. Muramatsu and Kawai's survey [2] emphasizes real-time anomaly detection in big data, aligning with the project's focus on processing high-frequency energy data in real time. Shankar and Ranjan [3] explored machine learning for anomaly detection in smart grids, guiding the use of models like Isolation Forest for industrial applications. Ahmed et al.'s survey [4] deepened the understanding of diverse techniques, refining the model selection. Lee and Zhang [5] discussed anomaly detection in industrial IoT systems, highlighting scalability and adaptability in handling dynamic data, key objectives for this project. By integrating these insights, this project aims to develop a user-friendly, real-time anomaly detection system with interactive chatbot support, empowering industries to improve efficiency, minimize downtime, and promote sustainability.

II. LITERATURE SURVEY

Anomaly detection has become a vital tool in the industrial sector, especially in energy-intensive industries like steel manufacturing, where it helps identify inefficiencies, reduce energy consumption, and minimize operational costs. With the growing complexity and volume of industrial data, traditional anomaly detection methods are increasingly inadequate for handling high-dimensional, dynamic datasets. This literature survey focuses on advancements in anomaly detection techniques, particularly those using machine learning, which are applicable to the steel manufacturing industry and similar energy-intensive sectors.

1. Anomaly Detection Using Self-Organizing Maps (SOM)

Kwon and Han (2008) [1] introduced a novel approach to anomaly detection using self-organizing maps (SOM). Their work highlighted the potential of neural networks in identifying irregular patterns in large datasets. SOM, a type of unsupervised learning algorithm, is capable of reducing high-dimensional data into simpler, more manageable forms

while retaining essential features for anomaly detection. The ability of SOM to cluster similar data points together allows for the identification of outliers or anomalies by recognizing data that deviates significantly from the typical patterns. This study significantly influenced the development of machine learning-based anomaly detection frameworks, leading to the integration of neural networks like SOM in detecting anomalies within complex industrial datasets, such as energy consumption data in steel manufacturing.

2. Real-Time Anomaly Detection in Big Data

The work of Muramatsu and Kawai (2018) [2] emphasizes the importance of real-time anomaly detection in big data systems, especially in the context of industrial environments where decisions need to be made quickly. Real-time processing is crucial in systems like energy monitoring in steel manufacturing, where the timely identification of irregularities can prevent inefficiencies, equipment failures, and unsafe conditions. Their survey reviewed various real-time anomaly detection techniques, including clustering, statistical models, and machine learning algorithms, for big data systems. They also highlighted challenges in processing high-frequency data, a key concern in industries like steel manufacturing where operational data can be highly dynamic and high-volume. This review aligns with the current project's emphasis on designing a real-time anomaly detection system that processes high-frequency energy consumption data, ensuring that anomalies are detected and flagged promptly to enable proactive decision-making.

3. Machine Learning Applications in Energy Monitoring

Shankar and Ranjan (2021) [3] explored the applications of machine learning techniques for anomaly detection in smart grid systems. Smart grids, which involve the integration of sensors, communication networks, and intelligent computing for efficient energy management, have much in common with energy monitoring in industries like steel manufacturing. Their work focused on the application of algorithms such as support vector machines (SVM), k-means clustering, and isolation forests to detect irregularities in energy usage patterns. Isolation Forest, a particularly efficient algorithm for detecting anomalies in high-dimensional data, was highlighted as a robust technique due to its ability to handle large datasets with minimal computational cost. This insight influenced the choice of the Isolation Forest algorithm in the current project, which aims to detect anomalies in energy usage and CO₂ emissions in real-time within steel manufacturing processes.

4. Comparative Survey of Anomaly Detection Techniques

Ahmed et al. (2016) [4] provided a comprehensive survey on network anomaly detection techniques, which, while focusing on cybersecurity, also included several general methods applicable to industrial anomaly detection. Their survey categorized methods into statistical models, machine learning algorithms, and deep learning approaches. The comparison between these methods revealed the strengths and weaknesses of each, particularly in terms of their ability to handle high-dimensional and dynamic data. Statistical methods, while effective for simpler, smaller datasets, often struggle with the complexity of real-world industrial data. Machine learning methods, such as decision trees and ensemble techniques, are more capable of adapting to changes in data distribution, making them well-suited for energy consumption data in steel manufacturing. Ahmed et al.'s findings were crucial in selecting the appropriate machine learning models for the current project, which aims to combine robustness and scalability for detecting anomalies in energy-intensive industrial settings.

5. Anomaly Detection in Industrial IoT Systems

Lee and Zhang (2017) [5] focused on anomaly detection in industrial Internet of Things (IoT) systems, which are becoming increasingly prevalent in sectors like steel manufacturing. IoT devices, such as sensors, collect massive amounts of real-time operational data, including energy consumption, temperatures, equipment status, and emissions. Lee and Zhang's study emphasized the importance of scalability and adaptability in anomaly detection models to handle the dynamic nature of IoT-generated data. They reviewed various approaches, including statistical methods, machine learning algorithms, and hybrid techniques, discussing their application in fault detection, predictive maintenance, and energy optimization. Their research directly aligns with the objectives of the current project, which aims to develop a scalable, real-time anomaly detection system for industrial IoT systems, ensuring efficient monitoring and reducing operational costs in steel manufacturing.

6. Integration of Machine Learning with Real-Time Monitoring Systems

The combination of machine learning techniques with real-time monitoring systems has shown significant potential in improving operational efficiency and sustainability in various industries. Several studies highlight the application of machine learning models like decision trees, random forests, and support vector machines in real-time anomaly

detection, enabling systems to learn from past data and continuously adapt to new patterns. The integration of these models with graphical user interfaces and chatbots further enhances the usability of these systems, especially for non-technical stakeholders. By providing real-time notifications and actionable insights, these systems empower decision-makers to take immediate action and prevent costly failures. For example, in the steel manufacturing industry, real-time anomaly detection can identify irregular energy consumption patterns that indicate equipment malfunctions or inefficiencies, allowing operators to address issues before they lead to downtime or excessive energy use.

7. AI Chatbot Integration for Industrial Anomaly Monitoring

Recent advancements have shown the utility of AI-powered chatbots in industrial anomaly detection systems, enhancing human-machine interaction and operational transparency. One such example is the Bolt.New platform, which integrates a conversational AI chatbot that allows users to query system status, receive anomaly alerts, and obtain explanations in natural language. This improves system accessibility for non-technical users and ensures real-time decision-making support. The use of AI chat interfaces demonstrates the growing emphasis on intuitive user experiences in complex industrial environments, aligning with this project's goal of embedding a chatbot for improved usability and user engagement.

Reference	Objectives	Methods/Approach	Key Findings	Limitations
Kwon and Han (2008)	Introduce SOMs for anomaly detection in high-dimensional data	Self-Organizing Maps (SOMs)	Effective at pattern recognition and anomaly detection in large datasets	Scalability challenges for real-time applications
Muramatsu and Kawai (2018)	Explore real-time anomaly detection in big data systems	Survey of real-time detection methods for big data	Addressed challenges of high-frequency, large-scale data processing	Lacked practical implementation focus
Shankar and Ranjan (2021)	Investigate ML techniques for energy monitoring in smart grids	Isolation Forest and ensemble methods	Highlighted energy monitoring and fault detection algorithms suitable for industrial contexts	Limited adaptation for broader IoT applications
Ahmed et al. (2016)	Compare anomaly detection techniques across networks	Statistical, machine learning, and hybrid approaches	Comprehensive insights into various anomaly detection methods	Network-centric focus limits applicability to IoT or industrial systems
Lee and Zhang (2017)	Focus on industrial IoT anomaly detection	Scalable frameworks for dynamic IoT systems	Emphasized adaptability and scalability for dynamic industrial operations	Provided limited practical algorithmic implementation details
Qi et al. (2023)	Apply NLP for log-based anomaly detection	ChatGPT with chatbot integration for log analysis	Enhanced log anomaly detection and explanation using chatbot frameworks	Lacked validation for high-dimensional, numerical datasets
Wang et al. (2022)	Develop deep learning for time-series anomaly detection	CNN-LSTM-based deep learning framework	Superior performance in detecting temporal anomalies in sensor data	High computational requirements pose challenges for real-time deployment

III.EXISTING SYSTEM

Existing systems for anomaly detection in industrial environments employ various machine learning algorithms and traditional rule-based methods to identify operational risks and vulnerabilities. These systems primarily focus on detecting anomalies in structured datasets using algorithms like Support Vector Machines (SVMs), Random Forests, and Decision Trees. While effective in some cases, these methods often lack scalability for high-dimensional and real-time data, making them less adaptable to dynamic industrial environments.

Current solutions typically involve pre-collected datasets that are processed and used to train machine learning models. These models are tested for accuracy and then integrated into platforms with basic visualization and alerting

capabilities. While these systems provide a foundational approach to anomaly detection, they often rely on static datasets and predefined conditions, which can limit their effectiveness in handling evolving patterns in complex datasets.

While tools like Tesseract OCR and heuristic-based post-correction methods are used in some systems for text and anomaly recognition, the lack of real-time processing, intuitive interfaces, and advanced analytics reduces their usability in fast-paced industrial settings. These limitations highlight the need for a more robust and scalable solution that integrates advanced machine learning techniques, real-time analytics, and user-friendly interfaces to address the challenges of modern industrial environments effectively.

IV. PROPOSED SYSTEM

The proposed system leverages advanced machine learning techniques with TensorFlow and Flask to enhance real-time anomaly detection in industrial environments. The system introduces innovative approaches such as optimized data structures for standardized data storage, character encoding for converting data into a digital format, and robust data management techniques to minimize data loss. By employing models tailored to the dataset and refining them for optimal performance, the system ensures accurate and efficient detection of operational anomalies.

Built with modern technologies like Node.js, React, and Bolt. New, and deployed on Streamlit Cloud, the system features a user-friendly interface designed for seamless interaction. The Isolation Forest algorithm is central to the model, effectively detecting vulnerabilities and anomalies in high-dimensional industrial data. Data for training is pre-processed to ensure quality and relevance, collected from diverse sources, and used to train models capable of real-time analysis and prediction. The system evaluates model accuracy through continuous testing, ensuring reliable detection and response mechanisms.

The platform provides a comprehensive dashboard with real-time visual analytics and a chatbot for intuitive communication. Users can receive predictive insights and alerts on operational risks, empowering timely and informed decision-making. The solution is adaptable to various industrial use cases, enabling anomaly detection, risk mitigation, and efficient management of business-critical systems. By combining advanced algorithms with accessible tools, the proposed system addresses the challenges of industrial anomaly detection with precision and scalability.

V. CONCLUSION

The research highlights the crucial role of machine learning-based anomaly detection systems in improving operational efficiency and sustainability in energy-intensive industries like steel manufacturing. As industrial data becomes more complex, the need for real-time processing, scalability, and adaptability in detecting anomalies is critical. Machine learning techniques, such as self-organizing maps and Isolation Forest, have proven effective in identifying irregularities in high-dimensional datasets typical of industrial operations.

Our proposed system builds upon these techniques by leveraging advanced machine learning models, like Isolation Forest, combined with modern technologies such as TensorFlow, Flask, and Bolt. New, to provide a real-time anomaly detection solution for steel manufacturing. The system incorporates optimized data management techniques to handle high-frequency, dynamic data, ensuring accurate and efficient anomaly detection. A user-friendly graphical interface, along with an integrated chatbot, makes the system accessible to non-technical users, enabling real-time monitoring and decision-making.

While the system shows great promise, challenges such as processing high-frequency data, ensuring scalability, and maintaining model adaptability remain. Addressing these challenges will be essential for improving the system's performance and ensuring its relevance in evolving industrial environments.

In conclusion, the proposed system offers a scalable, real-time solution for energy monitoring and anomaly detection in steel manufacturing. Future work will focus on incorporating predictive analytics to detect anomalies proactively and enhancing the chatbot functionality with natural language processing techniques. As industrial IoT systems continue to evolve, ongoing research into adaptive anomaly detection will be key to addressing the growing complexity of industrial data.

REFERENCES

1. Kwon, S. J., & Han, K. (2008). Anomaly detection using a self-organizing map. *IEEE Transactions on Neural Networks*, 19(6), 946-956. doi:10.1109/TNN.2008.4781136
2. Muramatsu, M., & Kawai, H. (2018). Real-time anomaly detection for big data. *IEEE Access*, 6, 196-209. doi:10.1109/ACCESS.2017.2783240
3. Shankar, K., & Ranjan, A. (2021). Machine learning for anomaly detection in smart grids. *IEEE Access*, 9, 49309-49325. doi:10.1109/ACCESS.2021.3060367
4. Ahmed, M., Mahmood, A. N., & Hu, J. (2016). Survey of network anomaly detection techniques. *European Journal of Operational Research*, 254(2), 1037-1050. doi:10.1016/j.ejor.2015.04.023
5. Lee, J., & Zhang, Z. (2017). Anomaly detection in industrial IoT systems. *IEEE Internet of Things Journal*, 4(3), 750-760. doi:10.1109/JIOT.2017.2785882
6. Qi, J., Huang, S., Luan, Z., Fung, C., Yang, H., & Qian, D. (2023). LogGPT: Exploring ChatGPT for log-based anomaly detection. *arXiv preprint arXiv:2309.011*



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details