A Study on Outlier Detection in Smart Manufacturing Applications

Jeong-Hun Kim, Tserenpurev Chuluunsaikhan, Aziz Nasridinov*
*Dept. of Computer Science, Chungbuk National University
{etyanue, teo, aziz}@chungbuk.ac.kr
*Corresponding Author

Abstract

Smart manufacturing is a process of integrating computer-related technologies in production and by doing so, achieving more efficient production management. The recent development of supercomputers has led to the broad utilization of artificial intelligence (AI) and machine learning techniques useful in predicting specific patterns. Despite the usefulness of AI and machine learning techniques in smart manufacturing processes, there are many fundamental issues with the direct deployment of these technologies related to data management. In this paper, we focus on solving the outlier detection issue in smart manufacturing applications. More specifically, we apply a state-of-the-art outlier detection technique, called Elliptic Envelope, to detect anomalies in simulation-based collected data.

1. Introduction

Manufacturing is a driving force behind the economy of a country. With the emergence of the fourth industrial revolution, manufacturing industries have begun their transformation from manpower-oriented to the machine-oriented management approach. In other words, industries have switched to a smart manufacturing paradigm. Smart manufacturing is a process of integrating computer-related technologies in production and by doing so, achieving more efficient production management [1]. In this paper, we focus on smart manufacturing applications in the ceramic field, where various high-tech chips and devices are produced.

The main success behind smart manufacturing applications can be associated with an advance in computer technologies. The recent development of supercomputers has led to the broad utilization of artificial intelligence (AI) and machine learning techniques in various industries. In smart manufacturing applications, these technologies are useful in analyzing the complex manufacturing process and make valuable feedback to the factory workers in terms of predicting certain patterns. For example, we can use AI and machine learning techniques to predict machine break patterns or recommend a golden-recipe that can be used for the optimization of manufacturing processes.

Despite the usefulness of AI and machine learning techniques in smart manufacturing processes, there are many fundamental issues with the direct deployment of these technologies related to data management. In this paper, we focus on solving the outlier detection issue in smart manufacturing applications. Considering that smart manufacturing applications utilize a large amount of Internet of Things (IoT) devices, the data that comes from these devices must be first filtered out and then used for AI and machine learning analysis to avoid bias analysis. Outlier detection techniques can simplify this process by identifying suspicious data points that can cause bias analysis. In this paper, we apply a state-of-the-art outlier detection technique, called Elliptic Envelope, to detect anomalies in simulation-based collected data. We also demonstrate how the outlier detection phase can operate with dimensionality reduction using state-of-the-art correlation analysis.

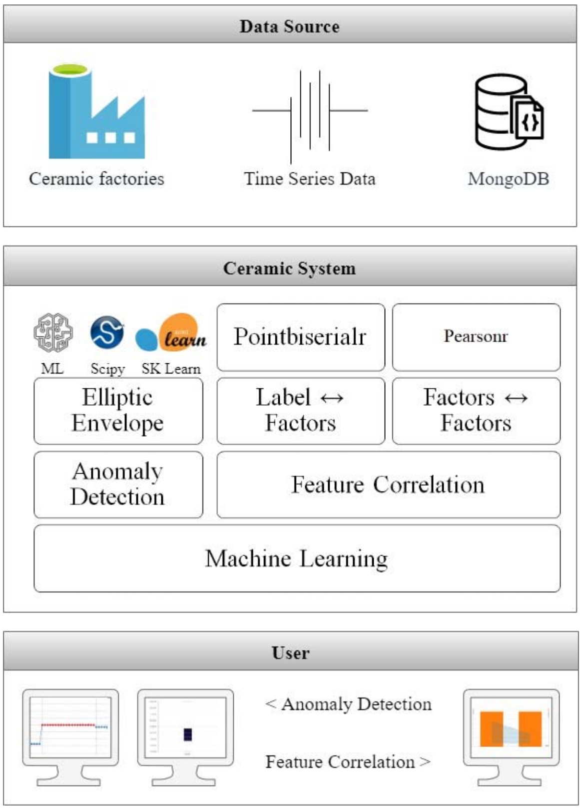
2. Related Study

In this chapter, we describe the related study. Chan et al. [2] proposed to modeling manufacturing processes using a genetic programming-based fuzzy regression. The proposed method takes advantage of a combination of two methods: genetic programming and fuzzy regression to demonstrate high accuracy in outlier detection. Gharibnezhad et al. [3] proposed a new method for detecting damages in manufacturing products. For that, the authors propose a robust Principle Component Analysis (PCA) algorithm that can detect damages in the presence of outliers. The proposed method was tested in an aircraft turbine blade for detecting the damages early in the production process. Nahar et al. [4] proposed a statistical approach for detecting outliers and determining screening parameters for efficiently identifying the outliers in manufacturing processes. The proposed method is utilized for production wafer probe data to identify outliers, i.e. at-risk material.

3. Proposed System

In this chapter, we describe the overall proposed system in detail. The ceramic is a web-based system shown in Figure 1. Our system intends to help to make a data-driven decision to factories that produce ceramic products. To that, we analyze big data related to factory manufacturing and visualize it. We divide the explanation of the proposed system into three parts: data collection, data analysis, and data visualization.

The data collection part is responsible for gathering data from various sensors in the manufacturing environment. It is important to note that it is difficult to obtain manufacturing process-related data from ceramic producing factories because of the data is highly secure. To avoid this situation, we collected the data from a simulation environment modeled specifically for the ceramic producing factory, which is stored in the MongoDB database. Specifically, we received 315 MongoDB documents that account for 281 Kb of data. The data contains the fields related to the local factory in South Korea, such as factory name, line, process, machine, product name, production data, and product specification.

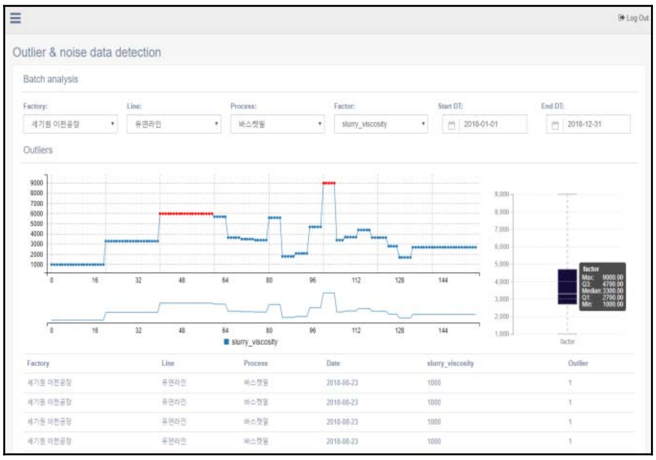


(Figure 1) Outlier detection process in smart manufacturing

The data analysis part is responsible for preprocessing the data and detecting outliers. Anomaly detection is a process to obtain data points that deviate from normal or expected. EllipticEnvelope is one of the anomaly detection algorithms which uses unsupervised or supervised. The algorithm fits Gaussian distribution to the dataset. Feature correlation is one of the processes in machine learning. Based on the process, we select the most relevant features to apply machine learning algorithms. Features selection can use feature correlation as label to factors, and factors to factors. In the label to factors, all numerical factors are compared to a classification factor. In the factors to factors, all numerical factors are compared to other numerical factors. We use Point biserial, Pearson algorithms to the label to factors, the factors to factors feature correlation.

4. System Implementation

In this chapter, we describe the system implantation, which is the third part of the proposed system. We use JavaScript open-source libraries like Jquery, D3, and others in front-end, and the Django framework of Python is used for the back-end of the system. Figure 2 demonstrates a user-interface for detecting outliers in smart manufacturing processes. The user-interface contains three parts. The first part shows the system configuration, such as manufacturing factory, line, dates, and others. The system configuration enables the user to have flexible data navigation. The second part demonstrates the visualization of time-series data that incorporates real-time outlier detection. Here, the red data points are outliers. The third part shows the input data in the form of a table.



(Figure 2) User-interface of outlier detection in smart manufacturing application

5. Conclusion

In this short paper, we have demonstrated a system that can detect outliers in the ceramic manufacturing industry. The proposed system offers a user-interface that enables factories to monitor an industrial process, collect historical data, and predict action in the future. In the future, we are planning to experimentally demonstrate the accuracy of various outlier detection methods as well as adding more data processing modules to the proposed system.

Acknowledgement

The work is supported by the Ministry of Trade, Industry & Energy (MOTIE, Korea) under i-Ceramic manufacturing innovation platform technology development business. No.2004367, 'Development of cloud big data platform for the innovative manufacturing in ceramic industry'.

References

[1] F. Tao, Y. Tang, X. Zou, Q. Qi, "A field programmable gate array implemented fibre channel switch for big data communication towards smart manufacturing," Robotics and Computer-Integrated Manufacturing, Vol. 57, June 2019, pp. 166-181.

[2] K. Y. Chan, C. K. Kwong, T. C. Fogarty, "Modeling manufacturing processes using a genetic programming-based fuzzy regression with detection of outliers," Information Sciences, Vol. 180, No. 4, 15 February 2010, pp. 506-518.

[3] F. Gharibnezhad, L. E. Mujica, J. Rodellar, "Applying robust variant of Principal Component Analysis as a damage detector in the presence of outliers," Mechanical Systems and Signal Processing, Vol. 50–51, January 2015, pp. 467-479.

[4] A, Nahar, K. M. Butler, J. M. Carulli, C. Weinberger, "Quality improvement and cost reduction using statistical outlier methods," 2009 IEEE International Conference on Computer Design, pp. 64-69, February 2010, Lake Tahoe, CA, USA.