Name: Koloina RAKOTOBE

Name of your Level 1: Sarah JAADAN

Paper title: Performance Analysis of IoT-Based Sensor, Big Data Processing, and Machine Learning Model for Real-Time Monitoring System in Automotive Manufacturing

Source: Google Scholar Paper Link

Keywords specific to the paper: monitoring system; IoT-based sensor; big data processing; fault detection; DBSCAN; Random Forest

Summary of the main contributions:

This article revolves around an advanced system that is designed to help managers highly and effectively monitor their business processes. It is important to assess the enormous increased volume of data, particularly in manufacturing processes and monitoring systems for business process management lifecycles. The paper accentuates on the development of a real-time monitoring system for the automotive manufacturing process utilizing (Internet of Things) IoT-based sensors, big data processing, and a hybrid prediction model. For further explanation, the study elaborates on the characteristics of an IoT-generated sensor data, the system's architecture, implementation details, and performance metrics for the functioning and monitoring of the hybrid AI model. It showed that integrating IoT-based sensors with a big data processing system to monitor the Korean company's system (case study in the article) is effective for processing and analyzing large amounts of sensor data in real-time. It also provided a solution for further studies regarding the model implementation as it successfully collected and transmitted the data within seconds with low computational cost. The big data processing system developed in this study utilizes multiple software applications for experiments to support the model. The results demonstrate that the system is utilizable and can process continuous sensor data more efficiently than traditional models.

Additionally, fault and errors detection are something that often occurs in the manufacturing process, detecting or preventing them can show and predict whether the process is functioning normally or not. They propose a hybrid prediction model that is composed of (Density-Based Spatial Clustering of Applications with Noise) DBSCAN-based outlier detection and (Random Forest) RF classification. These are AI algorithms models that help with assessing, monitoring and controlling data. The results showed that the proposed hybrid prediction model shows efficient monitoring capabilities and improved fault prediction accuracy compared to the other models tested for the automotive manufacturing process industry. Furthermore, security is a big issue when more IoT devices are adopted, implemented, and connected in this system. That is why the paper proposes that the security of IoT devices and platforms should be considered in a future study. In addition, it is said that a variety of abnormal conditions during the manufacturing process should continue to be identified and collected so the proposed hybrid prediction model in the paper can be utilized by companies and keep on improving when it comes to predictions. Regardless, the article provides a critical and interesting view into the integration of IoT-based sensors, big data processing, and a hybrid prediction model for real-time monitoring in the manufacturing industry. This paper is relevant for BPM in a company because it offers a comprehensive approach and optimization when it comes to process monitoring and controlling. As for management, it would enable better decision-making and prevent unexpected losses caused by faults during their manufacturing process. These are big factors to consider in companies.

AI model used

- **Introducing the AI models**

AI models are composed of algorithms or computational systems that can perform tasks like humans with advanced intelligence, just like the one proposed in the article. In fact,

this article talks about a very innovative model as it integrates multiple techniques/algorithms to showcase the results and based on the researchers' propositions. It is a predictive hybrid model in the monitoring phase of a BPM lifecycle. It can analyze data, learn from it, and make predictions or decisions based on the patterns it discovers with time during the automotive manufacturing process.

In this article, to improve prediction accuracy they combined DBSCAN, Random Forest with IoT-based sensors data.

DBSCAN is an unsupervised machine learning model used for clustering data points. It means the model utilizes data to separate outliers from normal sensor data. Random Forest is a popular machine learning algorithm used to model classification and regression tasks. This AI model is used to predict faults given the IoT-based sensors data as an input. IoT-based sensors are physical devices equipped with sensors and communication capabilities that gather data from the surrounding environment and transmit it over the internet or other networks. These sensors can collect many types of data, such as temperature, humidity, motion, light, sound, or location information. The data gathered by IoT sensors can be processed through many software applications, to be used for monitoring, analysis, automation, and decision-making during the business process.

- **How do they contribute the idea proposed by the paper?**

The IoT-based sensors provide a continuous stream of data, which is a key characteristic for the hybrid model to analyze and make predictions in real-time. The data is processed on a platform/a remote server every 5 seconds, allowing for the continuous flow of data necessary for the hybrid model's fault detection and prediction capabilities, while utilizing software applications to assess its velocity and volume. DBSCAN is utilized to separate outliers from sensor data, while RF is employed to predict faults given the sensor data as input. Then, the results are presented to the manager in real-time via a web-based monitoring system. This combination allows for a more nuanced and accurate identification of abnormal events during the manufacturing process. For instance, when there is an anomaly detected, applying DBSCAN on IoT sensor data permits the identification of malfunctions in the monitored process. RF model can be trained to predict future states or events based on sensor data, like equipment failures or maintenance. Decision support and optimization are sure to be enhanced with this complex combination of models.

The creation of this hybrid prediction model demonstrates that it is more effective than a simple model for monitoring the automotive manufacturing process because it provides a more comprehensive and extensive approach to fault detection and using multiple algorithms can increase the results. This combination provides a robust and reliable solution for real-time monitoring and fault prevention in the automotive manufacturing industry. The hybrid model indeed provides real-time data collection, effective data processing and fault predictions monitoring. The article mentions that several AI classification models such as Naïve Bayes (NB), Logistic Regression (LR), Multilayer Perceptron (MLP), and Random Forest (RF) were tried alongside our hybrid model. But the hybrid model remains the most effective one as it achieved 100% in terms of accuracy when it comes to detection, monitoring and controlling.

Supported by a software application?

Since our hybrid model utilizes IoT-based sensors to collect data, its big data processing and monitoring systems are using Apache Kafka, Apache Storm, and MongoDB.

As for the IoT-based sensor device, the study points out that it is based on Raspberry Pi to collect data, which is a small-size, low-cost and powerful computer device.

NoSQL MongoDB is important because it stores data from the manufacturing process. Since it is an automotive manufacturing system, it needs to be managed by a software app that can offer scalability, as we use very big volumes of data daily.

Additionally, Apache Kafka is used as a message queue to handle the big amounts of sensor data collected during the manufacturing process through a platform based on OSS (open source software) that is cost-effective for implementation and integration.

Apache Storm uses a real-time processing engine that analyzes the flows of data from Apache Kafka and performs real-time prediction using the hybrid prediction model. It shows that it can identify whether the manufacturing process is functioning normally or not.

These applications all play their roles in supporting the new hybrid model by providing a strong infrastructure. Apache Kafka efficiently handles the data streams, Apache Storm performs real-time processing and prediction, and NoSQL MongoDB stores the sensor data. This approach gives the system an effective monitoring of the manufacturing process, permits us to identify faults, and improves decision-making for management.