IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 19 April 2023, accepted 8 May 2023, date of publication 17 May 2023, date of current version 24 May 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3276942

RESEARCH ARTICLE

Critical Data Detection for Dynamically Adjustable Product Quality in IIoT-Enabled Manufacturing

SACHIN K. SEN[1,2], (Member, IEEE), GOUR C. KARMAKAR[1], (Member, IEEE), AND SHAONING PANG[1], (Senior Member, IEEE)

[1]Institute of Innovation, Science and Sustainability, Federation University Australia, Ballarat, VIC 3350, Australia
[2]School of Computing, Electrical and Applied Technology, Unitec Institute of Technology, Auckland 1025, New Zealand

Corresponding author: Sachin K. Sen (sachin_sen@hotmail.com)

This work was supported by the Defence Science Institute (DSI), Australia, through the DSI-Facilitated Project.

ABSTRACT The IIoT technologies, due to the widespread use of sensors, generate massive data that are key in providing innovative and efficient industrial management, operation, and product quality control processes. The significance of data has prompted relevant research communities and application developers how to harness the values of these data in secure manufacturing. Critical data analysis, identification of critical factors to improve the manufacturing process and critical data associated with product quality have been investigated in the current literature. However, the current works on product quality control are mainly based on static data analysis, where data may change, but there is no way to adjust them dynamically. Thus, they are not applicable for product quality control, at which point their adjustment is instantly required. However, many manufacturing systems exist, like beverages and food, where ingredients must be adjusted instantaneously to maintain product quality. To address this research gap, we introduce a method that identifies the critical data based on their ranking by exploiting three criticality assessment criteria that capture the instantaneous product quality change during manufacturing. These three criteria are – (1) correlation, (2) percentage quality change and (3) sensitivity for the assessment of data criticality. The product quality is estimated using polynomial regression (POLY), SVM, and DNN. The proposed method is validated using wine manufacturing data. Our proposed method accurately identifies critical data, where SVM produces the lowest average production quality prediction error (10.40%) compared with that of POLY (11%) and DNN (14.40%).

INDEX TERMS IIoT-enabled manufacturing, data criticality, criteria, product quality control, wine-quality.

I. INTRODUCTION

The Industrial Internet of Things (IIoT) is an integration of the Internet of Things (IoT), intelligent computing, and big data along with information technology (IT) and operation technology (OT). IIoT leverages modern industrial applications, and production processes in manufacturing [1]. IIoT applications range from large national infrastructures such as power and chemicals to food & beverage industries [2]. An intelligent wine manufacturing system is a potential application that uses IIoT technologies to control wine quality using sensor data. IIoT-based digital technologies are fuelling manufacturing and production systems more than ever. Therefore, the efficiency and productivity of manufacturing systems have increased to meet high expectations [3] and compete with other industries. Subsequently, digital innovations will advance the manufacturing process for secure production quality.

Wine is one of the highly consumable alcoholic beverages, the demand for which grew worldwide, even during the Covid-19 pandemic situation [4]. The digitization of wine manufacturing using IIoT technology would significantly

The associate editor coordinating the review of this manuscript and approving it for publication was Nitin Nitin.

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License.
For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/

contribute to its production process for secure product quality. The factors that impact the quality of products are significant for product quality to gain more market space. IoT sensors in the IIoT instigated technology advantages in collecting sensor data to analyze and detect data criticality. Due to the digitization of industries, a substantial amount of data are available related to the manufacturing and company core processes towards model-based analysis [5].

Data-driven decision-making is growing more than ever due to the enormous amount of data generated in the intelligent modern industries, which need to be processed professionally [6]. It is a crucial and core challenge of managing data quality towards context-specific utilization and decision making [7]. Critical data has been analyzed for many IIoT application areas. Some examples of these include - (1) analysis of data criticality for operation & production management [8], (2) manufacturing performance improvement, and (3) improving the overall productivity of the manufacturing systems [9]. Shin et al. in [10] critically analyzed product usage data and developed a method for decision support to address product degradation status. Liu et al. [11] proposed a system criticality analysis method to support safety-critical complex systems by analyzing their reliability. Gupta and Mishra analyzed the criticality of a conventional milling machine to address the failure effect for selecting appropriate maintenance strategies [12].

For predictive maintenance decision support, a critical analysis of computer-based numerical lathe subsystems control has been studied in [13]. The critical data analysis in those works mainly focuses on improving the manufacturing processes and production management systems, which add value to the intelligent technology-based manufacturing arena. The main shortcoming of these critical analyses is that they are mainly static value analyses of recorded data for production management and operation.

Static value analysis is to analyse a component or part of a product before its application; an example is a connecting rod to move from a vehicle crankshaft's piston rotating activity, where the design and material of the rod need to be analysed before applying to the engine [14]. Besides, dynamic value analysis has been well-studied in biomedicine [15], statistics [16], [17], and other fields [18], [19]. Dynamic value analysis uses the characteristics of ingredient components during the process execution, and the components are adjusted instantaneously. Here, the components instantaneously adjusted means the adjustment of the amount of that component which is coupled with other components that exist at a particular time during the manufacturing process. For example, malware detection has become challenging due to its rapid growth, and adversaries are devising new techniques to evade detection methods. Dynamic analysis of malware behaviour results in more accurate detection than the static method, which uses pre-defined signature-based detection [19]. Similarly, product quality needs to be optimised in food and beverage by making instant adjustments to its ingredient components. For example, in the case of cooking sauce, if it tastes bitter, it needs an adjustment instantly with sour and sweet. In the fast-growing use of sensors in industrial process monitoring, data-driven information and process values analyses have been studied widely and applied over the past decades [20]. The dynamic value analysis of these sensor data ensures product quality through instantaneous monitoring to face the challenge of achieving goodwill in the market.

In addition, product quality and production costs are also required in the modern world to meet the increased global competition [21]. Therefore, the demand for product quality improvement is increasing to address this global competition in the industrial sectors. As alluded before, many manufacturing systems, such as food and beverage, must adjust their components dynamically to control the product quality. It is necessary to adjust the product component(s) dynamically to improve the quality of many products. Because changes in the number of individual components or characteristic values subsequently cause changes in the quality of the product. Precisely when the quality of the product depends on the individual component values.

Apart from operation and management, manufacturing companies also critically monitor product quality. These critical monitoring systems become essential when the product quality varies dynamically with the amount of individual ingredient or component values. A possible example could be the quality of wine; wine quality depends on its characteristics or component values. The characteristics of wine quality, such as acidity, density, taste, smell, and colour, are the deciding factors for the quality of the wine. Those are the possible reasons for monitoring and analyzing the components. The effectiveness of these monitoring systems depends on the accurate and instantaneous detection of critical data related to the ingredients. Here, data represents the value of ingredients over time and indicates the criticality of a particular ingredient at a specific time.

Product quality indicators are a range of values. For wine, product quality is represented between 1 and 10. There is also a reference wine quality value (example 7.0). If the components change the product quality value for the reference quality value, the component impacts the quality. Product quality is measured by its taste, colour and so on. In order to ensure the quality of wine, it requires instantaneous dynamic adjustment of components, where those are co-related. Product components become critical when they affect the product quality significantly. For example, suppose the quality is affected by the excess amount of one component at a particular production stage. In that case, the product quality needs to be adjusted by adding an appropriate amount of another component that negates the impact on product quality. In the case of producing wine, if pH becomes higher, it needs to increase the acidic components, such as volatile acidity being the critical component at that production stage.

However, there exist approaches for critical data analysis [22], identification of critical factors for improving management [23] and operation [24] of manufacturing plants, and critical data detection associated with product quality management & monitoring [25], [26]. To our knowledge, the existing works on critical data detection for product quality are based on static value analysis. This issue motivates us to develop a method to detect critical data affecting product quality and to control the product quality by dynamically adjusting a set of mutually coupled data. However, to our knowledge, such dynamic data value analysis in identifying mutually coupled critical sensory data does not exist in the current literature. The significant contributions of this research project are as follows:

1. In this paper, for the first time, a theoretical model to detect the critical data acquired from the sensors of an IIoT-based manufacturing system has been introduced.
2. Leveraging or developing the following data criticality assessment criteria that are the bedrock of the proposed model mentioned above.
   (a) Correlation measures considering the impact on an individual component of the product to the respective estimated product quality.
   (b) Percentage (%) of product quality change for a particular component of the product.
   (c) Sensitivity of the product quality for an individual product component.
3. The proposed model is developed based on the predicted instantaneous product quality values. These predictions are performed using non-machine learning and machine learning techniques such as polynomial regression, SVM, and DNN.
4. The performance and validation of the proposed model are conducted using wine manufacturing data available in [27]. The experimental results show the accurate identification of critical data. In contrast, the SVM-based prediction model shows fewer average prediction errors (10.40%) than the other two models (11% for POLY and 14.4% for DNN).

Even though the efficacy of our critical data detection is assessed using the wine dataset, applying the proposed approach has a broader scope, where the product quality can be improved by adjusting the ingredient components. Examples of product manufacturing include food & beverages and other manufacturing industries, such as steel manufacturing, where components are required to adjust dynamically to control the quality of the product. The method developed in this research can serve as a template for deploying IIoT devices to gather additional data to help detect product quality.

Notations and symbols used in this report are stipulated in TABLE 1.

The rest of this article is structured as follows: Section II discusses IIoT-based manufacturing preliminaries, which include IIoT-based manufacturing processes, intelligent wine manufacturing and its challenges. Section III discusses this research related to currently available works focusing on critical components in the manufacturing process, existing methods of critical data detection, as well as intelligent wine manufacturing and possible challenges, while Section IV proposes a critical data detection method and that includes critical data detection criteria and critical data detection methodology. Section V presents the experiments and results, while its subsections reflect the experimental setup and dataset, results and findings, measured the performance, compare different criteria adopted and provide an evaluation of experiments. Finally, Section VI concludes this research article by outlining a summary of the outcomes of the research contribution.

TABLE 1. Symbols and notations with nomenclature.

Symbol/ Notation	Nomenclature & Brief Description	Definition of Notations/Symbols
ν^i	Wine Characteristics (independent variables)	ν^i contributes to control wine quality, $where\ i \in (1, 2, \cdots, n)$.
Q_j	Reference wine quality	Q_j is the actual wine quality for sample size $j \in (1, 2, 3, \cdots, m)$.
Q_j^i	Estimated wine quality and is a dependent variable	Q_j^i is the estimated product quality of the i^{th} product component for the j^{th} data instance.
ϱ	Correlation co-efficient	Measures correlations between wine characteristics and quality.
$P_i^{\prime}$	Percentage of wine quality changes	Quantify the impact of characteristics on wine quality based on % of changes.
ξ_i	Sensitivity	Sensitivity of changes in wine characteristics values to that of the estimated wine quality.
FA	Fixed Acidity	The fixed acidity for example tartaric acid is measured as g/dm^3
VA	Volatile Acidity	Volatile acidity shows the amount of acetic acid in g/dm^3
CA	Citric Acid	Citric acid comes from grape berries and is measured as g/dm^3
RS	Residual Sugar	The measurement of residual sugar shows the sugar content in wine as g/dm^3
CL	Chlorides	Amount of chlorides is controlled using Sodium Chloride (NaCl) in g/dm^3
FSD	Free Sulfur Dioxide	FSD in the wine is maintained as mg/dm^3
TSD	Total Sulfur Dioxide	In wine, TSD also is controlled as mg/dm^3
DT	Density	Density of wine is measured in g/dm^3
PH	pH	Measures the strength of the acidic component, affects wine colour, its stability and oxidation. pH scale in wine ranges from 3 to 4.
SLP	Sulphates	Sulfur compounds such as Potassium Sulphate (K_2SO_4), g/dm^3
ALC	Alcohol	Amount of alcohol content in wine measured as % vol.

II. IIoT-ENABLED MANUFACTURING PRELIMINARIES

A digital revolution has occurred in the industry with the adoption of IIoT in manufacturing systems. Manufacturing execution systems (MES) are designed to analyze IoT sensor production data. The sensing of non-contact architectural acquisition collects big data from the production process that influences the product quality in the digital manufacturing systems [28]. In this context, product quality control is conducted to large extent through analyzing those functional data [29], [30], [31], [32], [33].

A. IIoT-ENABLED MANUFACTURING PROCESS

IIoT leverages digital manufacturing with the combination of IT and OT, on which physical objects along with the manufacturing infrastructures are networked with wireless sensors [34]. Compared to traditional manufacturing, IIoT-enabled digital manufacturing utilizes various types of sensors ubiquitously in monitoring the performance of equipment and machinery, managing the production process, and even controlling the product quality [35]. This advance in manufacturing has changed the work process of different company sections, ranging from people management and systems maintenance to product quality testing/control.

As such, enormous data growth in digital manufacturing enables decision-making more efficiently than ever, resulting in a faster production process, which not only expands the production capacity but also saves costs and energy [36]. In the end, this data growth will create manufacturing intelligence across the factory through the sensor-driven real-time data analysis and communication within the production environment [37]. As we know, the decision-making of traditional manufacturing in both developed and developing countries is mainly static and primarily based on staff experiences. The IIoT-enabled manufacturing has a digitized production process, which generates a large volume of data and makes decision-making more reliant on data analytics [38].

For intelligent manufacturing, the data collected are used to model dynamical functions of components and factors for improving production processes [39]. Due to the data growth in IIoT-enabled digital manufacturing technologies, it has become necessary to analyze the criticality of product components for leveraging data-driven decision-making to develop predictive modelling and control the product quality [40]. Data-driven decision-making has become the highest preference for controlling the manufacturing process proactively to ensure better product quality in IIoT-enabled digital manufacturing [41]. In the manufacturing industry processes such as chemical or steel production, where components for production require to adjust instantaneously, data analysis for the criticality detection is a priority for controlling the production process [41].

In the food and beverage industry, such as digitalized wine manufacturing, the production control process needs to be dynamically adjusted considering the variation of crucial wine characteristics. Therefore, critically analyzed data will anchorage the production of high-quality wine.

B. IIoT-ENABLED DIGITAL WINE MANUFACTURING

The traditional wine manufacturing process encompasses steps ranging from grape crushing, pressing, fermentation, and clarification to ageing and bottling. In the traditional process, wine manufacturing is not automatic, production stages are primarily human-controlled, and the wine quality, in the end, is determined by staff experiences [42].

With IIoT implementation, digital transformations are conducted over the wine manufacturing process. As a result, sensors are deployed on a large scale to collect and validate sensory data for production process management. For product quality control, wine characteristics such as aroma, sweetness, bitterness, acidity, colour, and appearance [43] are under monitoring as those measurements are evolving at different stages of the production process.

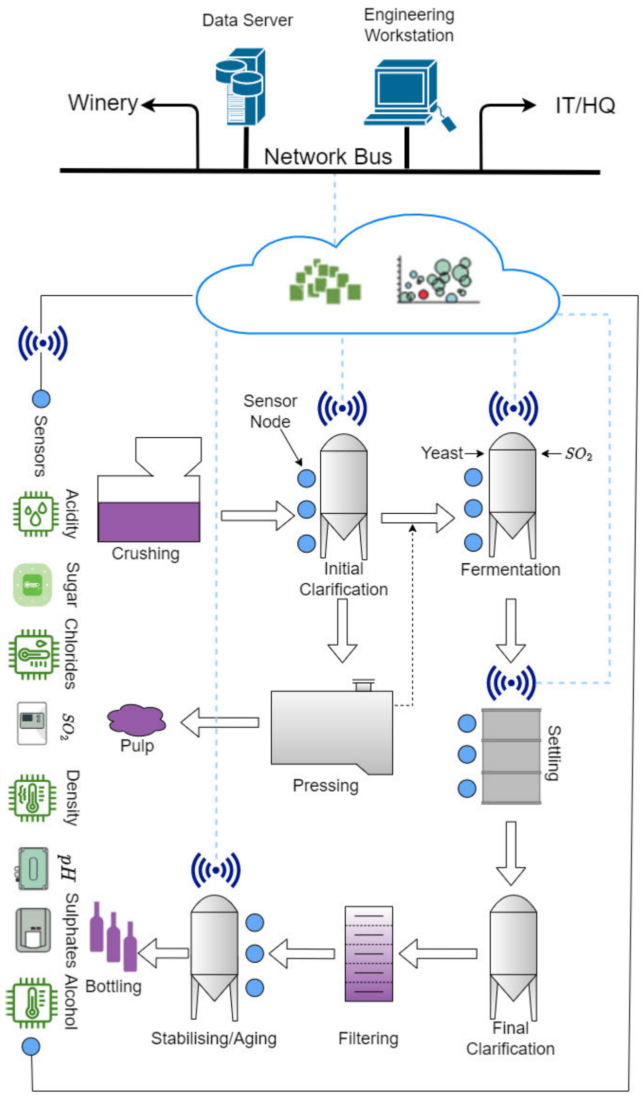

FIGURE 1. An illustration of IIoT-enabled wine manufacturing process, where data sensing is conducted in the respective process of different production stages, and captured data are transferred to the data server and processed in a separate engineering workstation.

For example, fermentation converts grape juices to wine using yeast, glucose & fructose, and other components like SO_2. This stage produces ethanol and controls the alcohol volume with the concentration of ethanol. Also, fermentation manages the wine aroma in that sugar content increases with fully ripe grapes, and less riped grapes result in acidity, and the increased pH value [44], [45]. For wine quality control, the above wine characteristics data need to be analyzed comprehensively to develop a generic working model that assists in computing their impact on the wine quality.

FIGURE 1 presents an IIoT-enabled digitized wine manufacturing process that represents different stages of wine production. The stages include initial clarification, fermentation, wine settling in cask/tanks, final clarification, filtering, and bottling of the product after stabilizing/ageing. As seen from the figure, during the manufacturing process, the deployed large-scale IIoT sensors monitor the measurements of wine components at different stages of production. The data are used to adjust wine components instantaneously to control the quality of the wine.

Sensors are placed in the different stages of the wine manufacturing system; measured sensor data are used to monitor the quality of the wine. Therefore, these sensor data play a vital role in ensuring wine quality. The sensors used in monitoring wine quality are acidity, sugar, chlorides, SO_2, density, pH, sulphates, and alcohol sensors. For example, one of the sensors, like METTLER TOLEDO's InLab Max Pro-ISM sensor, is used to measure the pH of wine. Other sensors, such as optical sensors, are used to determine the SO_2, MQ3 alcohol sensor for measuring the level of alcohol, and DLO-M2_ex density sensor for measuring density in wine. The influences of some of these sensor data on the wine quality are explained in Section III-C. 'Acidity' sensors are responsible for measuring 'Fixed Acidity', 'Volatile Acidity', and 'Citric Acid'. As shown in TABLE 2, all three acidity characteristics correlate with wine quality, and the correlation values are 0.11, -0.38 and 0.21, respectively, revealing their impact on the quality of the wine.

Similarly, the sensors that measure sugar, chlorides, SO_2, density, pH, sulphates, and alcohol, are also quite influential while ensuring the wine quality. TABLE 2 shows that alcohol, sulphate, and volatile acidity have more impact on wine quality than others. Their correlation values are 0.48, 0.38 and -0.38, respectively.

III. RELATED WORKS

This section presents a review on data criticality detection with a focus on quality production in IIoT-based manufacturing systems. The review addresses firstly the critical components of general digital manufacturing, followed by existing methods for critical data detection in the manufacturing process, and the critical components for wine manufacturing.

A. CRITICAL COMPONENTS OF DIGITAL MANUFACTURING

In manufacturing processes, components including data, factors, and characteristics that may cause a deficit in productivity and business outcome would be considered critical. Those components play an essential role in product quality control. Compromised application of the components would cause a catastrophic impact on the production outcome. The success of IIoT-based smart manufacturing relies on how smartly the industry can manage and analyse data. Critical data monitoring enhances decision-making smartly [46] for improving productivity and product quality. Therefore, input and output are necessary to monitor and detect defects in each production phase. For maintaining or improving a product's quality, many products whose component(s) need to be adjusted dynamically.

For example, the pH value is such a dynamic component of cheese that it affects the ageing process of Mozzarella. It influences the functional characteristics, including differences in calcium and moisture contents [47]. Therefore, the amount of pH needs to be controlled by adding or reducing components (e.g., acidic components) that affect the pH values. In another example, components such as Ni-Cr alloys, iron-based powder, Ni-based wasp-alloy, and low-carbon steel powder are used [48] in manufacturing corrosion-resistant steel. These components can significantly impact even minor changes of any of those for the product quality [48]. Thus, in practice, it is necessary to carefully manipulate the components that control the effect of another component(s) to secure good quality production.

B. EXISTING METHODS OF DETECTING DATA CRITICALITY

Data criticality detection has previously been studied for production process optimisation. For criticality assessment of plant productivity in mechanical industries, Jasiulewicz-Kaczmarek et al. [49] analysed production management, production method planning and equipment maintenance and machinery. They aimed at improving and maintaining the productivity of manufacturing plants. However, they considered the criticality assessment criteria of a machine and the interactions between them. This criticality assessment focuses only on the machine and the devices but not the critical factors associated with the product ingredients. Similarly, Zheng et al. [50] analyses the production-dependent data for improving the efficiency of the machinery production process. Antosz and Ratnayake [51] analysed the classification of machinery and their prioritisation for effectively scheduling preventive maintenance. These factors support decision-making for maintaining and improving plant productivity, again, not securing product quality.

Maintenance is another critical factor for productivity. Marquez et al. [52], and Singh et al. [53] analysed engineering and operational data related to asset functional loss and its frequency of occurrence to prioritise management of maintenance works. In [54], Stadnicka et al. investigated machine failures & costs due to the failures, machine up & downtime as well as deterioration of product quality in order to address environmental challenges, as well as mitigation of health and safety issues. Bengtsson in [55] performed a classification of management and technical personnel, including production management, production and maintenance teams, to prioritise improvement activities and maintenance programmes according to different machine types. Authors in [56] assessed engineering management data for prioritising assets (e.g., equipment and tools) to meet the business targets by aligning actions for maintenance. Moore and Starr [57] prioritised industrial maintenance jobs based on plant conditions where they used a strategy on cost criticality, making

TABLE 2. Correlations between wine characteristics and their estimated wine quality.

Wine Characteristics and Estimated Quality	Fixed Acidity	Volatile Acidity	Citric Acid	Residual Sugar	Chlorides	Free SO_2	Total SO_2	Density	pH	Sulphates	Alcohol	Wine Quality
Fixed Acidity	1	-0.28	0.66	0.22	0.25	-0.18	-0.09	0.62	-0.71	0.21	-0.70	0.11
Volatile Acidity	-0.28	1	-0.61	0.03	0.16	0.02	0.09	0.03	0.23	-0.33	-0.22	-0.38
Citric Acid	0.66	-0.61	1	0.18	0.11	-0.08	0.01	0.35	-0.55	0.33	0.10	0.21
Residual Sugar	0.22	0.03	0.18	1	0.21	0.07	0.15	0.42	-0.09	0.04	0.12	0.03
Chlorides	0.25	0.11	0.16	0.21	1	0	0.13	0.41	-0.23	0.02	-0.28	-0.19
Free SO_2	-0.18	0.02	-0.08	0.07	0	1	0.79	-0.04	0.12	0.05	-0.08	-0.06
Total SO_2	-0.09	0.09	0.01	0.15	0.13	0.79	1	-0.13	-0.01	0	-0.26	-0.20
Density	0.62	0.03	0.35	0.42	0.41	-0.04	0.13	1	-0.31	0.16	-0.46	-0.18
pH	-0.71	0.23	-0.55	-0.09	-0.23	0.12	-0.01	-0.31	1	-0.08	0.18	-0.04
Sulphates	0.21	-0.33	0.33	0.04	0.02	0.05	0	0.16	-0.08	1	0.21	0.38
Alcohol	-0.07	-0.22	0.10	0.12	-0.28	-0.08	-0.26	-0.46	0.18	0.21	1	0.48
Wine Quality	0.11	-0.38	0.21	0.03	-0.19	-0.06	-0.20	-0.18	-0.04	0.38	0.48	1

a trade-off between risks and costs. To prioritise spare parts planning and maintenance scheduling, the research in [58] examined the day-to-day priorities of the workforce and the machinery settings.

In summary, most existing data criticality detection research is from the mechanical industry for productivity improvement in operation, production management and maintenance, in which maintenance and workforce management are given priority. All these works have significantly contributed to expediting manufacturing processes' productivity. However, to our knowledge, these methods do not focus on identifying critical dynamic value data associated with product characteristics that affect product quality.

C. CRITICAL DATA COMPONENT IN WINE MANUFACTURING

Wine manufacturing, like other industries, involves the production management of equipment/machinery, operation, maintenance, and personnel management. It is worth noting that in an IIoT-enabled digital manufacturing platform, the quality of wine, which accomplishes from its colour, flavour, taste and smell, relies solely on how we capture those factors and perform sensor data detection. Technically, the sensory data of Fixed Acidity, Volatile Acidity, Citric Acid, Residual Sugar, Chlorides, Free Sulpher Dioxide, Total Sulphur Dioxide, Density, pH, Sulphates, and Alcohol control the quality of the wine while manufacturing. These components are interrelated among themselves, as can be seen in TABLE 2. Due to mutual coupling, some of those components significantly contribute to the quality of the wine. A careful analysis of their characteristics and respective criticality is necessary to quantify the amount of impact on the wine quality. The organic acids in grapes and the acid produced from alcoholic fermentation influence the acidity in wines [59]. The organic acid of grapes contains citric acid, malic acid and tartaric acid, which contribute to titratable acidity in the wine significantly [59]. Volatile acidity, total acidity, and alcohol impact the quality of wine; volatile acid at a higher level of pH affects wine quality significantly [60].

The acidity components impact wine taste by inversely affecting a potential wine component pH because if acidity increases, pH decreases. The pH is an influential component that has an equal or higher impact on wine quality, can affect the taste and colour and could cause cloudiness in the wine due to an increase of iron phosphate [61]. The pH plays a vital role in oxidation phases. The advanced oxidation processes of winery wastewater largely depend on pH [62]. SO_2 is the component that improves the fermentation process and prevents secondary fermentation [63]. This component reduces acidity to some extent, and an increase of SO_2 will cause a decrease in alcohol in wine. Also, SO_2 helps maintain the quality of wine [63] for colour stabilization and inhibition of biogenic amine synthesis.

The pH is inversely related to the acidity (i.e., higher pH implies lower acidity and vice versa). The component influences the quality of wine together with other components such as volatile acidity, total acidity and alcohol. For example, volatile acid at a higher level pH worsens the taste of wine typically [60]. The pH value needs to be maintained between 3.2 and 4.0 to control the acidity and colour of the final product [64], [65]. During the wine manufacturing process, the pH values can be adjusted by adding or reducing the acidity contents since we know the increase of acidity values will bring down the pH.

Even though alcohol dominates the quality of wine, acidity components appear to be more critical for wine manufacturing. The criticality of a component varies over time and needs to be determined for a particular component (e.g., alcohol). The sugar components like residual sugar are produced during the grape berry cultivation, which impacts the density of wine [66]. The Chloride component controls the sodium chloride (NaCl) in the wine and positively correlates to the fixed acidity [58]. The correlation between the above wine components and their quality values is regarded as the data criticality for wine production. TABLE 2 presents the list of particular wine characteristics and their correlation with the estimated wine quality values in [67].

In securing product quality, those correlations are difficult to adjust appropriately because the correlation can not be measured quantitatively, and estimating the correlation requires considering the dynamic effect of multiple individual components.

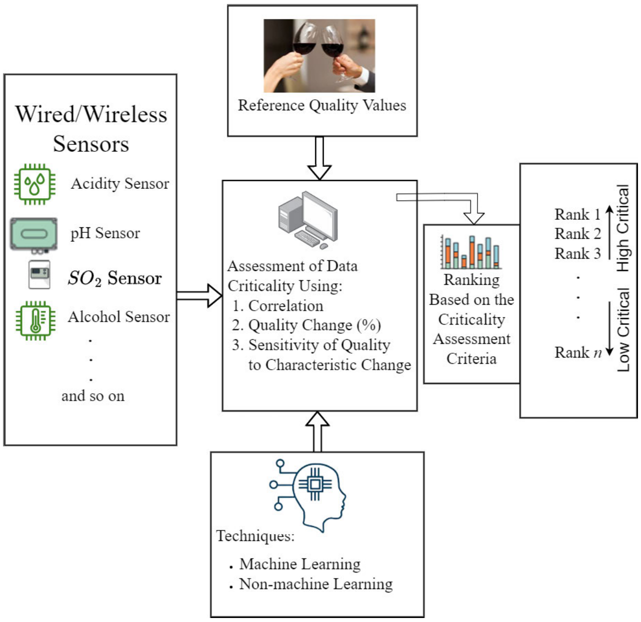

FIGURE 2. **Schematic diagram of how the proposed critical data detection method works. This diagram is developed considering the wine manufacturing context. However, this diagram applies to manufacturing processes where product quality changes dynamically.**

IV. PROPOSED CRITICAL DATA DETECTION METHOD

This section proposes a theoretical framework for detecting data criticality in IIoT-based intelligent manufacturing systems. The theoretical framework defines the data criticality detection criteria (metrics). And explains the details of assessing the criticality based on ranking.

A. OUTLINE OF THE PROPOSED METHOD

The key method is to assess the product quality change by computing the difference between predicted and reference quality for the product component value changes. FIGURE 2 presents a schematic diagram of how the proposed method uses three criticality detection criteria for detecting critical data. Three criteria include - (1) Correlation, (2) Percentage of quality changes, and (3) Sensitivity of components to product quality and applies to rank the criticality of data in an intelligent production system.

As stipulated from the schematic diagram in FIGURE 2, the proposed method comprises the aspects outlined below:

(1) The sensor data of a particular product characteristics are captured by the proposed method.
(2) The proposed method uses reference product quality. Note that this reference product quality is usually determined intuitively.
(3) The three critical data detection criteria to assess data criticality as mentioned above are leveraged in the method.
(4) Machine learning and non-machine learning modelling techniques are used to derive estimated product quality for each product characteristic. The non-machine learning technique may include polynomial regression (POLY), and machine learning techniques can be Support Vector Machine (SVM), deep learning algorithms like Deep Neural Network (DNN), or any suitable ones.
(5) The data criticality criteria such as correlation, percentage of product quality change, and sensitivity to product quality are calculated using the reference and estimated values for each product characteristic. As shown in FIGURE 2, the ranking of each of the product characteristics is determined using relevant data criticality value for each criterion.
(6) The rank numbers are equal to the number of product characteristics, $x \in \{1, 2, 3, \cdots n\}$, where '1' represents the highest rank and 'n' represents the lowest rank; higher the rank higher the criticality of data.

B. CRITICAL DATA DETECTION CRITERIA

For critical data detection, we use three criteria - (1) Correlation, (2) Percentage of quality change and (3) Sensitivity. These three criticality detection criteria are defined and discussed below:

1) CORRELATION BETWEEN PRODUCT COMPONENT AND ESTIMATED QUALITY

In practice, we calculate the correlation between i^{th} component values (ν^i) and their respective quality (Q^i); a suitable method is Spearman's rank correlation [68], [69] as,

$$\varrho = \frac{cov\left(\mathcal{R}(\nu^i), \mathcal{R}(Q^i)\right)}{\sigma_{\mathcal{R}(\nu^i)}\sigma_{\mathcal{R}(Q^i)}} = 1 - \frac{6\sum d_i^2}{n(n^2-1)}, \begin{cases} \text{when all } n \text{ rank values} \\ \text{are distinct integers} \end{cases} \quad (4.1)$$

where,

ϱ = Spearman's rank correlation coefficient
$cov(.)$ = Covariance between elements in the bracket
$\mathcal{R}(\nu^i)$ = Product component rank variables
$\mathcal{R}(Q^i)$ = Product quality rank variables
$\sigma_{\mathcal{R}(\nu^i)}$ = Standard deviation of rank variable, $\mathcal{R}(\nu^i)$
$\sigma_{\mathcal{R}(Q^i)}$ = Standard deviation of rank variable, $\mathcal{R}(Q^i)$
d_i = Difference between rank variables of product component and quality, $\left(\mathcal{R}(\nu^i) - \mathcal{R}(Q^i)\right)$
n = Number of rank variables

The correlation formula defined in Equation 4.1 represents linear relation between the product component and the reference quality. The correlation coefficients vary between 1 and −1. The positive and negative values indicate the percentage of increment and decrement for each other. The value of correlation coefficient 0 indicates no correlation, 1 indicates a perfect positive correlation, and that of −1 represents a perfectly negative correlation.

For example, for the wine manufacturing process, 11 characteristics (components are represented with characteristics in this particular case) are typically monitored to secure the quality of the wine. For the impact of a particular characteristic, i, its correlation between ν^i and Q^i will be

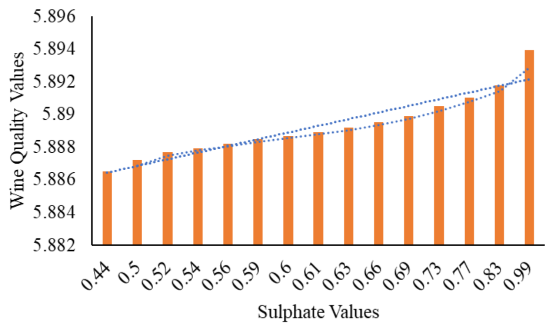

FIGURE 3. The correlation between sulphates (a product component) and the respective quality of the wine (product quality) has been presented. A trend-line indicates how the correlation exponentially varies between the component and the product quality values.

calculated without considering the impact of other characteristics.

FIGURE 3 gives an example of such a correlation between the sulphates values (ν^{10}) and the respective quality of wine Q^{10}. As we can see, with the increased sulphate values, respective wine quality increases, which is a non-linear increase. FIGURE 3 demonstrates that the relationship between the wine characteristics and its estimated wine quality is non-linear. However, in contrast, correlation represents a linear relationship vindicating the unsuitability of applying correlation in critical data detection. Therefore, it is necessary to develop different metrics to estimate the criticality of wine characteristics in terms of their impact on wine quality change. Two of such metrics are presented in the following section.

2) PERCENTAGE OF PRODUCT QUALITY CHANGE

For every component of the product and every data, we monitor the product quality change with respect to the reference product quality for that data. In doing that, for i^{th} component, we calculate the quality change per j^{th} value of the quality instances as,

$$\Delta_i = \sum_{j=1}^{m} \frac{\left|Q_j^i - Q_j\right|}{Q_j} \tag{4.2}$$

where m represents the number of data instances or the quality values sample size for the i^{th} component. Q_j indicates the reference quality values of j^{th} data instance. Q_j^i is the estimated product quality of the i^{th} product component for the j^{th} data instance.

Note that the reference product quality for each data is the combined impact of all components. This combined impact necessitates considering their overall impact irrespective of whether individual component produces positive or negative impacts on product quality. Leveraging this principle, the percentage of the quality changes (P_i) for i^{th} the product components is defined as,

$$P_i = \frac{\Delta_i}{\sum_{i=1}^{n}\{\Delta_i\}} \times 100\% \tag{4.3}$$

where n is the number of components considered for the product quality measurement.

Equation 4.3 reflects the relative impact of product quality change for a particular component over the total quality change for all the components.

3) SENSITIVITY OF PRODUCT COMPONENTS TO THE QUALITY

Sensitivity analysis is a usual approach to evaluate how a production system is sensitive against changes in its control parameters [70]. Here, for a manufacturing plant, the sensitivity of individual components to quality change captures the impact of product quality change for that product component. Changes in the product components and subsequent changes in the product quality can be calculated using Equations 4.4 and 4.5 to compute the sensitivity (ξ_i) for the i^{th} product component using Equation 4.6 [71].

(1) Product quality change between the predicted and reference quality for the j^{th} data instance of the i^{th} component is $\Psi_j = |Q_j^i - Q_j|$. Total changes ($\Delta\Psi_j$) in the consecutive product quality value changes between j^{th} and $(j+1)^{th}$ data instances of a particular i^{th} component will be computed as,

$$\Delta\Psi_j = \sum_{j=1}^{n-1} |\partial\Psi_j| \tag{4.4}$$

where,

$\Delta\Psi_j$ = Total changes in the product quality
$|\partial\Psi_j|$ = $|\Psi_j - \Psi_{j+1}|$
j = Number of data instances (sample size)

(2) Total changes in the i^{th} product component ($\Delta\Phi_j$) in the product quality can be calculated as,

$$\Delta\Phi_j = \sum_{j=1}^{n-1} |\partial\Phi_j| \tag{4.5}$$

where,

$\Delta\Phi_j$ = Total changes in the product component
$|\partial\Phi_j|$ = $|\Phi_j - \Phi_{j+1}|$
j = Number of data instances (sample size)

Therefore, the sensitivity score ξ_i can be written as,

$$\xi_i = \frac{1}{n-1}\left[\frac{\Delta\Psi_j}{\Delta\Phi_j}\right] \tag{4.6}$$

Further to the above discussion, there are more advantages and disadvantages of the proposed criteria adopted, but not limited to the following.

(1) The correlation method measures the relationship between variables, which helps determine the possible

impact of an independent variable on the dependent variable. However, if multicollinearity exists among the variables, it could be problematic to determine the exact amount of impact of an independent variable on the dependent variable.

(2) The advantage of the percentage of quality change (PQC) is a measure of relative change for others which yields comparative criticality. For the components with different units of measurement, the loophole is that those units are not considered in the calculation.

(3) The advantage of sensitivity is that it directly reflects the impact of a particular component changes on the quality of the product. Therefore, sensitivity is not a relative measure like the PQC. The disadvantage is that even though other components' impact on the product quality exists, it does not account for that.

V. EXPERIMENTS AND RESULTS

We have used wine quality monitoring sensory data sourced from a database in [27] for experiments. The database contains both red wine and white wine datasets, but we considered the red wine dataset for experiment and analysis in this research. These data were collected from the production process of Portuguese Vinho wine.

The red wine dataset contains 11 independent wine characteristics and wine quality. The wine quality varies with the changes in individual characteristics of wine. The characteristics values, v^i and the quality value, Q represent the i^{th} wine characteristics and the wine quality respectively. Here, $i \in \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11\}$, which represents the index of the set {*"Fixed Acidity (FA)", "Volatile Acidity (VA)", "Citric Acid (CA)", "Residual Sugar (RS)", "Chlorides (CL)", "Free Sulpher Dioxide (FSD)", "Total Sulpher Dioxide (TSD)", "Density (DT)", "pH (PH)", "Sulphates (SLP)", "Alcohol (ALC)"*}.

A human testing process determines wine quality; therefore, this wine quality is referred to as reference wine quality in this paper. Each of the features (characteristics or quality) has 1,599 samples of data which were split into two parts: (*a*) training dataset and (*b*) test dataset. In the wine quality estimation process, capturing a wine characteristic's impact requires considering its various instantaneous values while keeping other characteristic values constant. This approach generated 1000 training data and 599 testing data samples for all eleven characteristics from the 1,599 data samples provided in [72].

A. EXPERIMENTAL SETUP

To assess the performance of different models, both non-machine learning and machine learning techniques - POLY, SVM and DNN were developed. Linear and POLY are two commonly used non-machine learning algorithms, of which we have used the higher degree polynomial regression due to the nonlinearity of our dataset. Similarly, varieties of machine-learning algorithms, where SVM and DNN are prevalent due to their better performance. As mentioned in [73], DNN and radial basis function (RBF) based SVM perform better than similar modelling algorithms regarding computational and prediction accuracy. We used the Python Scikit-Learn library to develop the models in our experiment. For DNN, one of the two architectures, the Multilayer Perceptron (MLP) model, was used and fed with eleven inputs to produce one output. The inputs include the wine characteristics FA, VA, CA, RS, CL, FSD, TSD, DT, PH, SLP, and ALC, where the output is the predicted wine quality. The performance of the DNN model was evaluated by computing Mean Square Error (MSE) and the Regression Scores (R2-Scores). The MSE and R2-Scores against different numbers of hidden layers are stipulated in FIGURE 4. In addition, we used the '*adam*' optimiser as the '*adam*' optimiser is widely used as a solver for weight optimisation. ReLU (rectified linear unit) was used as the activation function at the hidden layers. We can see from the plot in FIGURE 4 that the MSE reduces gradually with the hidden layer increase, and the MSE reaches the lowest value of 0.019503 for the hidden layer 3. On the other hand, respective R2-Scores remain almost the same for hidden layers 3 to 5, and the highest R2-Score, 0.971211, is achieved at hidden layer 3. Therefore, the model with 3 (three) hidden layers has achieved the highest precision level since it shows the lowest error and highest regression score. Therefore, we used three hidden layers in our DNN model for the experiment.

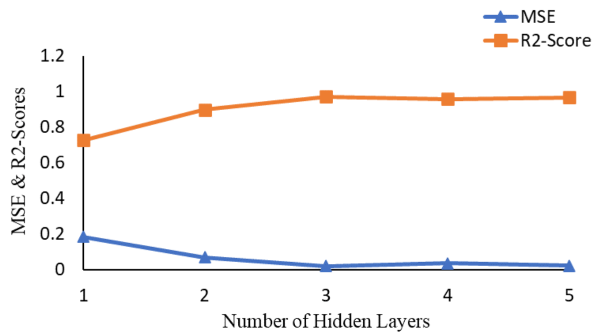

FIGURE 4. The MSE and the R2-Scores for different number of hidden layers of DNN.

We validated our DNN model using MSE as loss values as a performance indicator for a different number of epochs with a batch size of 200. The loss values are plotted against epoch numbers shown in FIGURE 5. FIGURE 5 shows that the loss value gradually decreases up to 500 epochs and then increases again. The model obtained the highest precision for 500 epochs with a minimum loss value of 0.010727, which motivated us to use 500 epochs in our experiment.

SVM uses several kernels, such as Linear, Polynomial, and RBF. For our experiment's better-performing kernel type, we evaluated the performance of those kernels by computing the MSE and R2-Scores during the training, which is plotted in FIGURE 6. FIGURE 6 shows that RBF performs better with the lowest MSE and highest R2-Score, 0.0296978 and 0.561627, respectively, indicating RBF kernel is the best one

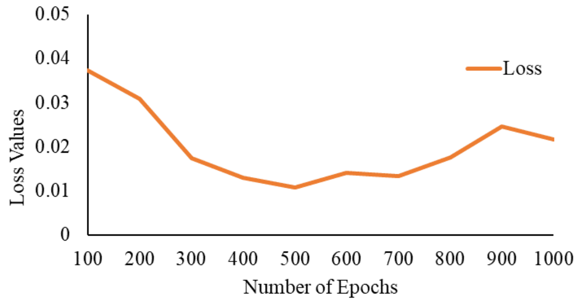

FIGURE 5. Loss values of the DNN training model for different number of epochs.

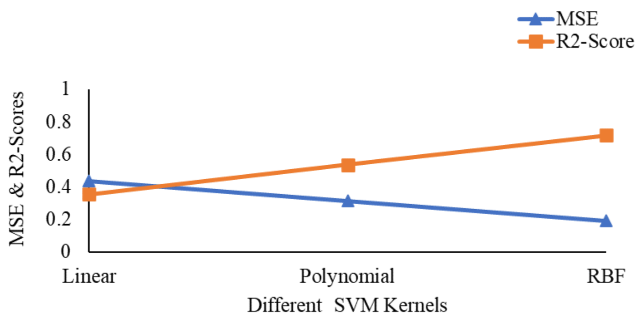

FIGURE 6. The MSE and the R2-Scores for different Kernel Types of SVM.

for our dataset. As the better performer compared to the other two kernels, we chose the RBF kernel to train the SVM model. The regularisation parameter, **C**, plays a significant role in the SVM with RBF kernel. The regularisation strength of the model optimisation is inversely proportional to the **C**-Value. Too small **C**-Value results in a useless model, and too high value obtains high estimation accuracy during training but low in the testing phase [74]; for the **C**-Value, we used **C**=1 in our experiment.

For the POLY model, we considered the higher degree polynomial regression as it performs better for non-linear systems. In our experiment, we used the 3^{rd} order polynomial as existing research shows that the 3^{rd} order polynomial produces the best fit of the data for most applications [75], [76]. The polynomial with more than 3^{rd} order is more likely to reach saturation and shows erratic behaviour [77].

Note that the MSE, R2-Score and loss values were calculated using validation dataset (20% of training data). Each of the model output was cross-validated with 10-fold and 100-run settings. All hyperparameter settings used in our experiment are presented in TABLE 3.

Using the wine quality values estimated by the three models mentioned above along with wine characteristics values, we have computed three criteria (metrics) for detecting data criticality (refer to the three data criticality assessment criteria defined in Section IV-B), which are,

(1) The correlation between the instantaneous data values of a wine characteristic and its estimated wine quality values.

TABLE 3. Hyperparameters and their respective values were used for the models POLY, SVM and DNN.

Model	Parameter	Value
POLY	Degree of polynomial	3
	n-splits	10
	n-repeats	3
	Random state	1
SVM	Kernel	RBF
	C (Regularisation parameter)	1.0
	Tol (Tolerance)	0.001
	Epsilon (loss function)	0.0
	Max iterations	1000
DNN	Number of hidden layers	3
	Number of neurons per layer	100
	Alpha	0.0001
	Tol (Tolerance)	0.0001
	Learning rate	0.001
	Epsilon (loss function)	1.0E-08
	Number of epochs	500

(2) The percentage of wine quality change for a particular wine characteristic.

(3) The sensitivity of wine quality for a wine characteristic, i.e., change in wine quality for a wine characteristic.

B. RESULTS AND DISCUSSION

In the wine quality estimation process, capturing a wine characteristic's impact requires considering its various instantaneous values while keeping other characteristic values constant. In this wine quality estimation process, we use the three models - POLY, SVM, and DNN. The output of the models is used to develop the critical data detection criteria values. The data criticality is ranked using each model's computed data criticality criteria values. The results (the values obtained from three criteria) are also used to assess the performance of all three data models to obtain the most suitable solution for the targeted scheme of the critical data detection method.

The following three subsections discuss the values of data criticality metrics developed for all 11 wine characteristics based on the results from all three models - POLY, SVM, and DNN. The values of three data criticality criteria metrics - correlation, wine quality change in % and sensitivity are shown in Figures 7, 8, and TABLE 4, respectively.

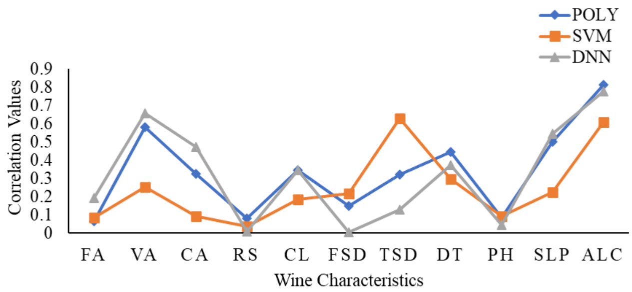

FIGURE 7. Correlations between the characteristics of wine and its respective estimated wine quality predicted by POLY, SVM and DNN.

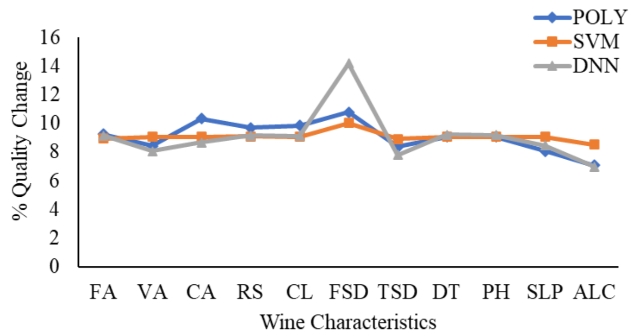

FIGURE 8. The PQC for wine characteristics resulting from the models - POLY, SVM and DNN.

1) CRITICALITY BASED ON CORRELATIONS

As explained in the critical data detection methodology, Section (IV-B), the correlation between wine characteristics and estimated wine quality is in [-1, 1]. Since both +ve and -ve correlation values represent the impact on wine quality, we have considered the absolute values of the correlation coefficients. We calculated the impact of wine characteristics on the quality of the wine based on the amount of impact only. FIGURE 7 shows the correlation values resulting from three models - POLY, SVM, and DNN. We compute the correlation-based criticality criteria from the data plots by assessing the amount of correlation.

The POLY model shows that the correlation value for ALC (0.8147) is the highest, followed by VA (0.5821), SLP(0.5005), and DT (0.4452). The rest of the correlation values are below 0.3, which can be considered insignificant. The prediction of SVM shows that the correlation value for TSD (0.6279) is the highest, followed by ALC (0.608). The remaining values are below 0.3, which also can be regarded as insignificant. The DNN model shows that the correlation value for ALC (0.7755) is at the top of this category. ALC is followed by VA (0.6562), SLP (0.5452), CA (0.4715), DT (0.3745) and CL (0.3437). The rest are less than 0.3, indicating that these characteristics are non-significant. The selection of correlation values below 0.3 to separate the unimportant characteristics is based on the fact that these values indicate weak correlation [78]; as alluded before, we have considered correlation values between 0 and 1.

The correlation measures show that the ALC values of POLY and DNN are highly significant for their respective categories; the VA values for DNN and POLY follow those correlation measures. For SVM, TSD is the highest, followed by ALC. Therefore, by applying the principle of our proposed method, that the higher the correlation, the higher the criticality is, we conclude that ALC, VA, and TSD are the most critical wine characteristics for the correlation criteria.

2) CRITICALITY BASED ON PERCENTAGE CHANGES OF WINE QUALITY

We compute the PQC as the ratio of the quality changes for each characteristic to the changes of all 11 characteristics. FIGURE 8 presents the PQC values for all 11 wine characteristics computed from the results of all three models - POLY, SVM and DNN.

The POLY model shows that the PQC for FSD (10.79%) is at the top of the list, followed by CA (10.33%) and CL (9.84%). The PQC for ALC (7.07%) is the lowest, and the remaining values vary from 8.0% to 9.0%. For SVM, the highest PQC is for FSD (10.04%), and the rest are in-between 9.11% and 9.21%, where ALC (8.51%) is an exception. For DNN, the PQC for FSD (14.18%) is on the top compared to the PQC values for TSD and ALC. The same as POLY, the remaining varies between 8.0% and 9.0%.

The PQC value for FSD is the highest for POLY and DNN in their respective category. The PQC values for CA from POLY and FSD from SVM follow the FSD values from the DNN and POLY. The predictions from POLY and SVM are very similar for most of the characteristics. The PQC of CA in POLY is the highest among others, almost the same as others in SVM, and is lower than five components but higher than the other five components in the DNN model. The trend of the contribution in ordered categories indicates that FSD contributes the highest amount, followed by CA. Therefore, according to the principle of our proposed method, the higher the PQC, the higher the criticality is, FSD and CA are substantially more critical than the rest.

3) CRITICALITY BASED ON SENSITIVITY

The sensitivities between the wine characteristics and their respective changes in wine qualities are shown in TABLE 4. TABLE 4 displays the sensitivity of wine quality changes for the changes in the wine characteristic values. The sensitivities are computed using the results obtained from the prediction models - POLY, SVM and DNN.

The results show that DT is the most sensitive among all the three models. The sensitivity values for DT are 25.5%, 29.37% and 32.52% for DNN, SVM and POLY, respectively. The CL, PH, CA, SLP, and VA sensitivities are almost in descending order for all three models. The corresponding sensitivity values of these wine characteristics are CL (1.09%, 1.15% and 1.72%), PH (0.28%, 0.47% and 0.41%), CA (0.30%, 0.28% and 0.41%), SLP (0.30%, 0.31% and 0.32%), VA (0.24%, 0.26% and 0.30%) for DNN, SVM and

TABLE 4. The sensitivity values of wine characteristics to the changes of estimated wine quality produced by the modeling techniques - POLY, SVM and DNN.

Sensitivity / Characteristics	Sensitivity Values		
	POLY	SVM	DNN
FA	0.0349	0.0279	0.0507
VA	0.2412	0.2593	0.3048
CA	0.3042	0.2831	0.4090
RS	0.0341	0.0257	0.0414
CL	1.0867	1.1524	1.7170
FSD	0.0077	0.0046	0.0062
TSD	0.0018	0.0016	0.0017
DT	25.518	29.374	32.524
PH	0.2843	0.4675	0.4059
SLP	0.3038	0.3108	0.3162
ALC	0.0407	0.0465	0.0678

POLY, respectively. The sensitivity values of the remaining characteristics (FA, RS, FSD, TSD, and ALC) are very low (less than 0.07%). The variations of the sensitivity values produced by all three models are minimum, which conforms to the consistent results produced by the three models for the sensitivity criteria.

The results predicted by the three models - DNN, SVM and POLY, show that the wine quality values are more-sensitive to DT, CL, PH, CA, SLP, and VA than others. Based on the principle "higher the sensitivity, higher the criticality is", as laid out in Section IV-B, these sensitivity values indicate that DT, CL, PH, CA, SLP, and VA are critical characteristics of the red wine, where DT is the most critical.

C. SUMMARY OF RESULTS

The values of criticality criteria for wine characteristics are ranked in TABLE 5. As shown in TABLE 5 and from the above results and subsequent discussions, different prediction models produce different impacts of wine characteristics on the quality of wine that are reflected in their ranking.

According to the critical data detection criteria outlined in Section IV-B, the higher the rank, the higher the criticality. The ranking values are between 1 and 11, where 1 is for the highest criticality and 11 for the lowest criticality. In TABLE 5 out of the 11 ranks, the top 6 (upper plane) ranked wine characteristics are categorised below for comparison among the criticality criteria developed from the results of our models - POLY, SVM and DNN. For POLY, the upper plane ranking of wine characteristics for three criticality criteria are as follows:

(1) ***Correlation***: VA - ***rank*** **1**, ALC - ***rank*** **2**, SLP - ***rank*** **3**, CA - ***rank*** **4**, CL - ***rank*** **5**, and DT - ***rank*** **6**.
(2) ***Percentage (%) of Quality Change***: PH - ***rank*** **1**, FSD - ***rank*** **2**, CA - ***rank*** **3**, FA - ***rank*** **4**, RS - ***rank*** **5**, and DT - ***rank*** **6**.
(3) ***Sensitivity***: DT - ***rank*** **1**, CL - ***rank*** **2**, CA - ***rank*** **3**, PH - ***rank*** **4**, SLP - ***rank*** **5**, and VA - ***rank*** **6**.

The POLY result reveals some commonality among the six critical wine characteristics. DT and CA are typical in all three criteria, CL, SLP and VA are common in the Correlation and Sensitivity. Between PQC and Sensitivity, the common characteristics are CA, DT and PH. For SVM, the upper plane ranking values are as follows:

(1) ***Correlation***: ALC - ***rank*** **1**, DT - ***rank*** **2**, PH - ***rank*** **3**, TSD -***rank*** **4**, FA - ***rank*** **5**, and CL - ***rank*** **6**.
(2) ***Percentage (%) of Quality Change***: FSD - ***rank*** **1**, RS - ***rank*** **2**, DT - ***rank*** **3**, CL - ***rank*** **4**, PH - ***rank*** **5**, and CA - ***rank*** **6**.
(3) ***Sensitivity***: DT - ***rank*** **1**, CL - ***rank*** **2**, PH -***rank*** **3**, SLP - ***rank*** **4**, CA - ***rank*** **5**, and VA - ***rank*** **6**.

For the results predicted by SVM, DT, PH and CL are typical in all the three critical data detection criteria, CA is only common for the PQC and Sensitivity criteria. For the DNN, the upper plane ranking values of wine characteristics are ranked as follows:

(1) ***Correlation***: ALC - ***rank*** **1**, VA - ***rank*** **2**, SLP - ***rank*** **3**, CA - ***rank*** **4**, DT - ***rank*** **5**, CL - ***rank*** **6**.
(2) ***Percentage (%) of Quality Change***: CA - ***rank*** **1**, DT - ***rank*** **2**, PH - ***rank*** **3**, CL - ***rank*** **4**, FSD - ***rank*** **5**, and RS - ***rank*** **6**.
(3) ***Sensitivity***: DT - ***rank*** **1**, CL - ***rank*** **2**, CA - **rank 3**, SLP - ***rank*** **4**, PH - ***rank*** **5**, and VA - ***rank*** **6**.

For the results obtained from DNN, CA, CL, & DT are common to all three criteria. VA, & SLP are common between Correlation and Sensitivity, where only PH is common between PQC and Sensitivity criteria.

For next level of assessment, we develop a Power Grid for the critical wine characteristics in the upper plane predicted by our models - POLY, SVM & DNN, and detected by the three criteria. The critical wine characteristics from the upper plane placed in Power Grid for assessment, are shown in TABLE 6.

As stipulated in TABLE 6, if we compare the elements of the power grid column-wise, we can see that DT & CA for the POLY, DT, CL & PH for SVM, and DT, CL & CA for the DNN model are common to all three criteria respectively. On the other hand, if we compare the power grid row-wise, it can be seen that only DT is common for all three models - POLY, SVM and DNN, irrespective of critical data detection criteria. For Correlation, DT, CL & ALC are common among POLY, SVM and DNN, VA, SLP & CA are common between POLY and DNN. For PQC, PH, FSD, CA, RS & DT are common among POLY, SVM, and DNN, only CL IS Common between SVM and DNN. For Sensitivity, all the wine characteristics for from each of the three models POLY, SVM, and DNN are common to each other respectively.

We can see that the wine characteristics appear multiple times in the power grid for different criteria (Correlation, PQC and Sensitivity) across the models POLY, SVM and DNN respectively. For example, the appearance frequencies of DT, CA/CL, PH, VA/SLP, ALC/FSD/RS, FA are 9, 8, 7, 5,

TABLE 5. The rank of wine characteristics produced by prediction models - POLY, SVM and DNN using the three critical data detection criterion: (1) Correlation, (2) % of Wine Quality Change, and (3) Sensitivity.

Characteristics \ Criteron/Model	Correlation			%Quality Change			Sensitivity		
	POLY	SVM	DNN	POLY	SVM	DNN	POLY	SVM	DNN
FA	7	5	7	4	8	7	8	8	8
VA	1	8	2	10	9	9	6	6	6
CA	4	9	4	3	6	1	3	5	3
RS	11	11	8	5	2	6	9	9	9
CL	5	6	6	8	4	4	2	2	2
FSD	10	10	10	2	1	5	10	10	10
TSD	8	4	9	7	11	11	11	11	11
DT	6	2	5	6	3	2	1	1	1
PH	9	3	11	1	5	3	4	3	5
SLP	3	7	3	9	7	8	5	4	4
ALC	2	1	1	11	10	10	7	7	7

TABLE 6. POWER GRID: Assessment of data criticality criteria as per the results obtained from the models - POLY, SVM and DNN.

Criterion \ Model	POLY	SVM	DNN
Correlation	VA	ALC	ALC
	ALC	DT	VA
	SLP	PH	SLP
	CA	TSD	CA
	CL	FA	DT
	DT	CL	CL
PQC	PH	FSD	CA
	FSD	RS	DT
	CA	DT	PH
	FA	*CL*	*CL*
	RS	PH	FSD
	DT	CA	RS
Sensitivity	DT	DT	DT
	CL	CL	CL
	CA	PH	CA
	PH	SLP	SLP
	SLP	CA	PH
	VA	VA	VA

3, and 2, respectively. Here, if we assume the principle that the higher the frequency of appearance, the higher the criticality is, DT, CA/CL, PH, VA/SLP, ALC/FSD/RS, FA are critical wine characteristics in the descending order.

However, since we consider the top six ranks, this principle treats a characteristic appears six times with rank 1 as the same as another characteristic appears six times with rank 6 in the worst case. Therefore, we need to calculate the rank by leveraging the rank value and the appearance frequency using the concept, weighted average method of categorisation. But, we can not use a weighted average using the frequency of appearance because the weighted average will not consider a characteristic appearing more frequently. This issue indicates that we need to include relative appearance frequency in the final ranking to ensure that the more frequently appearing characteristic possesses a higher rank. For that reason, we calculate the final rank termed normalized ranking, which is defined as,

$$\overline{R}_i = \frac{W_i}{A_i} \tag{5.1}$$

where,

$\overline{R}_i$ = Normalised weighted average of ranking
W_i = Weighted average of ranking
$= \frac{\sum_{i=1}(f_i \times R_i)}{\sum_{i=1} f_i}$
A_i = Appearance frequency ratio
$= \frac{f_a}{f_p}$
f_i = Appearance frequency
R_i = Rank of characteristics
f_a = Actual number of appearance frequency
f_p = Possible number of appearance frequency

The rank calculated using Equation 5.1 are shown in TABLE 7. TABLE 7 shows that DT, CA, PH, FSD, CL and SLP possess the higher six ranks, and the rank values are 1, 2, 3, 4, 5, and 6, respectively. These rank values are more analogous with the ranking produced by SVM predictions than those for DNN and POLY. The rationale behind the normalised weighted average adopted in this paper

TABLE 7. Normalised ranking using weighted average rank and appearance frequency.

Wine Characteristics	Weighted Average Ranking Value	Appearance Ratio	Normalised Ranking Value	Normalised Rank
FA	4.50	0.33	13.64	10
VA	4.20	0.50	8.40	8
CA	3.63	0.83	4.37	2
RS	4.33	0.50	8.66	9
CL	3.88	0.67	5.79	5
FSD	2.67	0.50	5.34	4
TSD	4.00	0.17	23.53	11
DT	2.90	1.00	2.90	1
PH	3.43	0.67	5.11	3
SLP	3.80	0.50	7.60	6
ALC	1.33	0.17	7.82	7

is that the weighted average ranking method is preferably more precise than the categorical process [79]. Moreover, in [80], the aggregated index score values for multi-objective decision-making are obtained considering weighted averages of respective parameters and the normalisation process is used to compare the criteria.

We conclude critical data detection methodology in another way by assessing the performance of the modelling techniques. For the performance assessment of our models - POLY, SVM, and DNN, we have performed an error calculation between the reference wine quality (Q_j) and estimated wine quality (Q_j^i) in the following section.

D. ERROR CALCULATION FOR ESTIMATED WINE QUALITY IN RESPECT OF REFERENCE VALUES

We perform the error calculation for estimated wine qualities for their respective reference values using widely used error calculation metrics, namely Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE). MAE performs better capturing model performance by reflecting the actual situation of the prediction error, and the RMSE measure the model prediction quality by selecting model error evaluation metrics [81].

1) RMSE FOR ESTIMATED WINE QUALITIES

In FIGURE 9, we have shown the differences between the average reference quality (Q) value and that of respective estimated quality (Q^i) values, which are computed using the prediction results of POLY, SVM and DNN.

There is a slight difference between errors produced by SVM and POLY. In contrast, the error difference between the SVM and DNN is quite significant, and the difference is consistent except for two closer points for TSD and the ALC. The DNN result has not overlapped at any point with SVM, but POLY overlaps in most points. Since the prediction error produced by SVM is the lowest for all 11 wine characteristics except ALC and VA, the data criticality ranking based on SVM is a better solution.

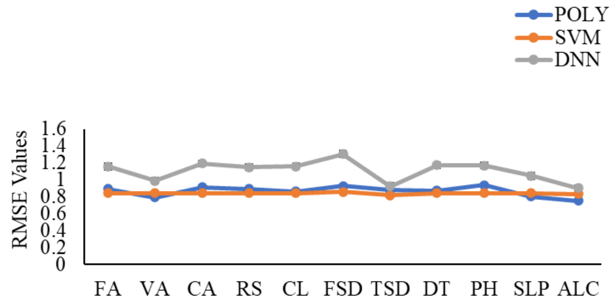

FIGURE 9. RMSE for estimated wine quality in respect of reference wine quality using POLY, SVM and DNN models.

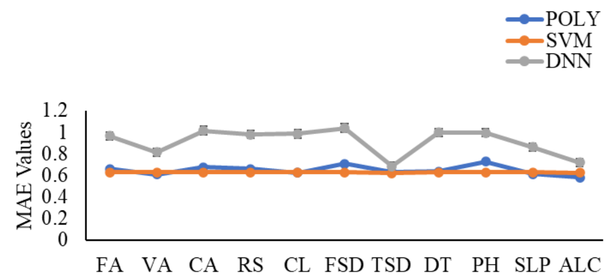

FIGURE 10. MAE for estimated wine quality in respect of reference wine quality using POLY, SVM and DNN based models.

We calculate MAE next to assess further the performance of the models - POLY, SVM, and DNN by capturing their prediction errors.

2) MAE FOR ESTIMATED WINE QUALITIES

We have also computed MAE between the average reference quality (Q) values and that of respective estimated quality (Q^i) values. FIGURE 10 shows the computed MAE results of POLY, SVM, and DNN models. The results of SVM are very similar to POLY, with slightly higher values for POLY at some points. However, the error values for DNN are significantly higher than those of SVM and POLY at most points.

There is a slight difference between errors produced by SVM and POLY for MAE and RMSE. In contrast, the error difference between the SVM and DNN is quite significant, and the difference is consistent except for two closer points for TSD and the ALC. The DNN result has not overlapped at any point with SVM, but POLY overlaps in most points. Since the prediction error produced by SVM is the lowest for all 11 wine characteristics except ALC and VA, the data criticality ranking based on SVM can be regarded as a better solution. Moreover, the average error values of both MAE (0.6266) and RMSE (0.8397) for the SVM model are lower than those of MAE (0.6480) and RMSE (0.8666) for POLY, respectively. The MAE and RMSE error values computed for DNN are higher, as seen in FIGURE10 and FIGURE9, respectively. Therefore, the SVM modelling technique would be a better solution.

VI. CONCLUSION

This paper's research work deals with a critical issue in the IIoT-enabled digital manufacturing process. Identifying critical data in the modern IT/OT integrated intelligent industrial production system is of utmost critical to protect secure production. This work addresses this crucial issue by introducing a method for identifying critical data for product quality control. We have analysed and assessed our method using a wine data set. In our method, we have estimated the data criticality using three assessment criteria - (1) correlation, (2) percentage of quality changes and (3) sensitivity. The estimated product quality for each wine characteristic is predicted using the models – POLY, SVM and DNN. We compute critical data using the three criteria, and subsequently, the most critical wine characteristics are determined by ranking them according to their influence to control the product quality. The ranking is influenced by the prediction error of POLY, SVM and DNN models. The model performance has been evaluated by prediction error computation using RMSE and MAE. The prediction error of the models is less than 14.5%, where SVM produces the minimum error (10.40%). This error indicates that the criticality ranking of wine characteristics has been performed reliably. Though the ranking is performed for all three criteria, the correlation is not a good choice because of its linear relationship between independent and dependent variables, which indicates that the other two criteria PQC and sensitivity, are more acceptable.

The conceptual framework developed in this work can be extended to other manufacturing areas, where product quality needs instantaneous adjustment of ingredient components. In future, we aim to evaluate the system using data from other relevant industries like food.

REFERENCES

[1] L. Peñalver and L. Parra, *Industrial IoT Technologies and Applications: 4th EAI International Conference, Industrial IoT 2020, Virtual Event, December 11, 2020: Proceedings*, vol. 365. Cham, Switzerland: Springer, 2021.

[2] K. Stouffer, V. Pillitteri, S. Lightman, M. Abrams, and A. Hahn, "Guide to industrial control systems (ICS) security," *NIST Special Publication*, vol. 800, p. 16, Jun. 2015.

[3] L. Monostori, "Cyber-physical systems in manufacturing," *CIRP Ann.*, vol. 65, no. 2, pp. 621–641, 2016.

[4] S. Colbert, C. Wilkinson, L. Thornton, and R. Richmond, "COVID-19 and alcohol in Australia: Industry changes and public health impacts," *Drug Alcohol Rev.*, vol. 39, no. 5, pp. 435–440, Jul. 2020.

[5] M. Schmidt, J. T. Maier, and L. Härtel, "Data based root cause analysis for improving logistic key performance indicators of a company's internal supply chain," *Proc. CIRP*, vol. 86, pp. 276–281, Jan. 2019.

[6] G. Schuh, E. Rebentisch, M. Riesener, T. Ipers, C. Tönnes, and M.-H. Jank, "Data quality program management for digital shadows of products," *Proc. CIRP*, vol. 86, pp. 43–48, Jan. 2019.

[7] T. Bauernhansl, S. Hartleif, and T. Felix, "The digital shadow of production—A concept for the effective and efficient information supply in dynamic industrial environments," *Proc. CIRP*, vol. 72, pp. 69–74, Jan. 2018.

[8] M. Gopalakrishnan, M. Subramaniyan, and A. Skoogh, "Data-driven machine criticality assessment—Maintenance decision support for increased productivity," *Prod. Planning Control*, vol. 33, no. 1, pp. 1–19, Jan. 2022.

[9] M. Gopalakrishnan, A. Skoogh, A. Salonen, and M. Asp, "Machine criticality assessment for productivity improvement: Smart maintenance decision support," *Int. J. Productiv. Perform. Manage.*, vol. 68, no. 5, pp. 858–878, Jun. 2019.

[10] J.-H. Shin, H.-B. Jun, C. Catteneo, D. Kiritsis, and P. Xirouchakis, "Degradation mode and criticality analysis based on product usage data," *Int. J. Adv. Manuf. Technol.*, vol. 78, nos. 9–12, pp. 1727–1742, Jun. 2015.

[11] L. Liu, D. Fan, Z. Wang, D. Yang, J. Cui, X. Ma, and Y. Ren, "Enhanced GO methodology to support failure mode, effects and criticality analysis," *J. Intell. Manuf.*, vol. 30, no. 3, pp. 1451–1468, Mar. 2019.

[12] G. Gupta and R. P. Mishra, "A failure mode effect and criticality analysis of conventional milling machine using fuzzy logic: Case study of RCM," *Qual. Rel. Eng. Int.*, vol. 33, no. 2, pp. 347–356, Mar. 2017.

[13] N. M. Thoppil, V. Vasu, and C. Rao, "On the criticality analysis of computer numerical control lathe subsystems for predictive maintenance," *Arabian J. Sci. Eng.*, vol. 45, pp. 5259–5271, Jul. 2020.

[14] A. Muhammad, I. H. Shanono, and I. Shanono, "Static analysis and optimization of a connecting rod," *Int. J. Eng. Technol. Sci.*, vol. 6, no. 1, pp. 24–40, 2019.

[15] W. Ye, G. Chen, X. Li, X. Lan, C. Ji, M. Hou, D. Zhang, G. Zeng, Y. Wang, C. Xu, W. Lu, R. Cui, Y. Cai, H. Huang, and L. Yang, "Dynamic changes of D-dimer and neutrophil-lymphocyte count ratio as prognostic biomarkers in COVID-19," *Respiratory Res.*, vol. 21, no. 1, pp. 1–7, Dec. 2020.

[16] A. Sharif, S. A. Raza, I. Ozturk, and S. Afshan, "The dynamic relationship of renewable and nonrenewable energy consumption with carbon emission: A global study with the application of heterogeneous panel estimations," *Renew. Energy*, vol. 133, pp. 685–691, Apr. 2019.

[17] J. N. Ng'ombe and B. W. Brorsen, "Bayesian optimal dynamic sampling procedures for on-farm field experimentation," *Precis. Agricult.*, vol. 23, pp. 2289–2311, Jun. 2022.

[18] A. Meixedo, J. Santos, D. Ribeiro, R. Calçada, and M. Todd, "Damage detection in railway bridges using traffic-induced dynamic responses," *Eng. Struct.*, vol. 238, Jul. 2021, Art. no. 112189.

[19] P. V. Shijo and A. Salim, "Integrated static and dynamic analysis for malware detection," *Proc. Comput. Sci.*, vol. 46, pp. 804–811, Jan. 2015.

[20] S. Zhang, C. Zhao, and B. Huang, "Simultaneous static and dynamic analysis for fine-scale identification of process operation statuses," *IEEE Trans. Ind. Informat.*, vol. 15, no. 9, pp. 5320–5329, Sep. 2019.

[21] H. Fatorachian and H. Kazemi, "A critical investigation of industry 4.0 in manufacturing: Theoretical operationalisation framework," *Prod. Planning Control*, vol. 29, no. 8, pp. 633–644, Jun. 2018.

[22] A. Alvanchi, A. TohidiFar, M. Mousavi, R. Azad, and S. Rokooei, "A critical study of the existing issues in manufacturing maintenance systems: Can BIM fill the gap?" *Comput. Ind.*, vol. 131, Oct. 2021, Art. no. 103484.

[23] S. Georgiev and S. Ohtaki, "Critical success factors for TQM implementation among manufacturing SMEs: Evidence from Japan," *Benchmarking, Int. J.*, vol. 27, no. 2, pp. 473–498, Mar. 2020.

[24] F. Psarommatis, P. A. Dreyfus, and D. Kiritsis, "The role of big data analytics in the context of modeling design and operation of manufacturing systems," in *Design and Operation of Production Networks for Mass Personalization in the Era of Cloud Technology*. Amsterdam, The Netherlands: Elsevier, 2022, pp. 243–275.

[25] C. Liu, L. Le Roux, C. Körner, O. Tabaste, F. Lacan, and S. Bigot, "Digital twin-enabled collaborative data management for metal additive manufacturing systems," *J. Manuf. Syst.*, vol. 62, pp. 857–874, Jan. 2022.

[26] M. S. Firouz, K. Mohi-Alden, and M. Omid, "A critical review on intelligent and active packaging in the food industry: Research and development," *Food Res. Int.*, vol. 141, Mar. 2021, Art. no. 110113.

[27] P. Cortez, A. Cerdeira, F. Almeida, T. Matos, and J. Reis, "Modeling wine preferences by data mining from physicochemical properties," *Decis. Support Syst.*, vol. 47, no. 4, pp. 547–553, Nov. 2009.

[28] X. Cai, L. Xue, and J. Cao, "Robust penalized M-estimation for function-on-function linear regression," *Stat*, vol. 10, no. 1, p. e390, Dec. 2021.

[29] J. O. Ramsay and C. J. Dalzell, "Some tools for functional data analysis," *J. Roy. Stat. Soc., Ser. B, Methodolog.*, vol. 53, no. 3, pp. 539–561, Jul. 1991.

[30] L. Horváth and P. Kokoszka, *Inference for Functional Data With Applications*, vol. 200. Cham, Switzerland: Springer, 2012.

[31] F. Ferraty and P. Vieu, *Nonparametric Functional Data Analysis: Theory and Practice*. Cham, Switzerland: Springer, 2006.

[32] P. Kokoszka and M. Reimherr, *Introduction to Functional Data Analysis*. Boca Raton, FL, USA: CRC Press, 2017.

[33] M. Colledani, T. Tolio, A. Fischer, B. Iung, G. Lanza, R. Schmitt, and J. Váncza, "Design and management of manufacturing systems for production quality," *CIRP Ann.*, vol. 63, no. 2, pp. 773–796, 2014.

[34] S. Jeschke, C. Brecher, T. Meisen, D. Özdemir, and T. Eschert, "Industrial Internet of Things and cyber manufacturing systems," in *Industrial Internet of Things*. Cham, Switzerland: Springer, 2017, pp. 3–19.

[35] D. P. Möller, "Digital manufacturing/industry 4.0," in *Guide to Computing Fundamentals in Cyber-Physical Systems*. Cham, Switzerland: Springer, 2016, pp. 307–375.

[36] T. Pereira, J. V. Kennedy, and J. Potgieter, "A comparison of traditional manufacturing vs additive manufacturing, the best method for the job," *Proc. Manuf.*, vol. 30, pp. 11–18, Jan. 2019.

[37] P. O'Donovan, K. Leahy, K. Bruton, and D. T. J. O'Sullivan, "An industrial big data pipeline for data-driven analytics maintenance applications in large-scale smart manufacturing facilities," *J. Big Data*, vol. 2, no. 1, pp. 1–26, Dec. 2015.

[38] F. Tao, Q. Qi, A. Liu, and A. Kusiak, "Data-driven smart manufacturing," *J. Manuf. Syst.*, vol. 48, pp. 157–169, Jul. 2018.

[39] J. M. Müller, O. Buliga, and K.-I. Voigt, "Fortune favors the prepared: How SMEs approach business model innovations in industry 4.0," *Technolog. Forecasting Social Change*, vol. 132, pp. 2–17, Jul. 2018.

[40] Q. P. He and J. Wang, "Statistical process monitoring as a big data analytics tool for smart manufacturing," *J. Process Control*, vol. 67, pp. 35–43, Jul. 2018.

[41] J. Krumeich, S. Jacobi, D. Werth, and P. Loos, "Big data analytics for predictive manufacturing control–A case study from process industry," in *Proc. IEEE Int. Congr. Big Data*, Jun. 2014, pp. 530–537.

[42] C. D. Porto, M. Cervellera, S. Galassi, and G. Lercker, "Carbonic maceration and traditional wine making: A quality control procedure using gas chromatographic determination of fixed compounds," *J. Wine Res.*, vol. 3, no. 2, pp. 97–104, Jan. 1992.

[43] S. Fuentes, E. Tongson, and C. G. Viejo, "Novel digital technologies implemented in sensory science and consumer perception," *Current Opinion Food Sci.*, vol. 41, pp. 99–106, Oct. 2021.

[44] T. Schöck and T. Becker, "Sensor array for the combined analysis of water–sugar–ethanol mixtures in yeast fermentations by ultrasound," *Food Control*, vol. 21, no. 4, pp. 362–369, Apr. 2010.

[45] R. S. Jackson, *Wine Science: Principles and Applications*. New York, NY, USA: Academic, 2008.

[46] M. Helu, K. Morris, K. Jung, K. Lyons, and S. Leong, "Identifying performance assurance challenges for smart manufacturing," *Manuf. Lett.*, vol. 6, pp. 1–4, Oct. 2015.

[47] M. A. S. Cortez, M. M. Furtado, M. L. Gigante, and P. S. Kindstedt, "Effect of pH on characteristics of low-moisture mozzarella cheese during refrigerated storage," *J. Food Sci.*, vol. 73, no. 9, pp. S443–S448, Nov. 2008.

[48] J. A. Cherry, H. M. Davies, S. Mehmood, N. P. Lavery, S. G. R. Brown, and J. Sienz, "Investigation into the effect of process parameters on microstructural and physical properties of 316L stainless steel parts by selective laser melting," *Int. J. Adv. Manuf. Technol.*, vol. 76, nos. 5–8, pp. 869–879, Feb. 2015.

[49] M. Jasiulewicz–Kaczmarek, K. Antosz, P. Żywica, D. Mazurkiewicz, B. Sun, and Y. Ren, "Framework of machine criticality assessment with criteria interactions," *Eksploatacja I Niezawodnoś́ Maintenance Rel.*, vol. 23, no. 2, pp. 207–220, Jun. 2021.

[50] L. Zheng, C. Zeng, L. Li, Y. Jiang, W. Xue, J. Li, C. Shen, W. Zhou, H. Li, L. Tang, T. Li, B. Duan, M. Lei, and P. Wang, "Applying data mining techniques to address critical process optimization needs in advanced manufacturing," in *Proc. 20th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2014, pp. 1739–1748.

[51] K. Antosz and R. M. C. Ratnayake, "Machinery classification and prioritization: Empirical models and AHP based approach for effective preventive maintenance," in *Proc. IEEE Int. Conf. Ind. Eng. Eng. Manage. (IEEM)*, Dec. 2016, pp. 1380–1386.

[52] A. C. Márquez, P. M. de León, A. S. Rosique, and J. F. G. Fernández, "Criticality analysis for maintenance purposes: A study for complex in-service engineering assets," *Qual. Rel. Eng. Int.*, vol. 32, no. 2, pp. 519–533, Mar. 2016.

[53] S. Singh, H. Singh, and R. Kumar, "Criticality analysis of equipments under maintenance policy by using multi criteria decision making," *Int. J. Appl. Eng. Res.*, vol. 10, no. 54, pp. 108–112, 2015.

[54] D. Stadnicka, K. Antosz, and R. M. C. Ratnayake, "Development of an empirical formula for machine classification: Prioritization of maintenance tasks," *Saf. Sci.*, vol. 63, pp. 34–41, Mar. 2014.

[55] M. Bengtsson, "Classification of machine equipment," *Age*, vol. 1, no. 2, p. 3, 2011.

[56] A. C. Márquez, P. M. D. León, J. G. Fernández, C. P. Márquez, and M. L. Campos, "The maintenance management framework: A practical view to maintenance management," *J. Maintenance Eng.*, vol. 15, no. 2, pp. 167–178, 2009.

[57] W. J. Moore and A. G. Starr, "An intelligent maintenance system for continuous cost-based prioritisation of maintenance activities," *Comput. Ind.*, vol. 57, no. 6, pp. 595–606, Aug. 2006.

[58] T. R. Moss and J. Woodhouse, "Criticality analysis revisited," *Qual. Rel. Eng. Int.*, vol. 15, no. 2, pp. 117–121, Mar. 1999.

[59] B. G. Defilippi, D. Manriquez, K. Luengwilai, and M. González-Agüero, "Aroma volatiles: Biosynthesis and mechanisms of modulation during fruit ripening," *Adv. Botanical Res.*, vol. 50, pp. 1–37, Jan. 2009.

[60] E. Mataix and M. D. L. de Castro, "Sequential determination of total and volatile acidity in wines based on a flow injection-pervaporation approach," *Analytica Chim. Acta*, vol. 381, no. 1, pp. 23–28, Feb. 1999.

[61] E. Peynaud and J. Blouin, *The Taste of Wine: The Art Science Wine Appreciation*. Hoboken, NJ, USA: Wiley, 1996.

[62] J. Rodríguez-Chueca, C. Amor, J. Mota, M. S. Lucas, and J. A. Peres, "Oxidation of winery wastewater by sulphate radicals: Catalytic and solar photocatalytic activations," *Environ. Sci. Pollut. Res.*, vol. 24, no. 28, pp. 22414–22426, Oct. 2017.

[63] R. F. Guerrero and E. Cantos-Villar, "Demonstrating the efficiency of sulphur dioxide replacements in wine: A parameter review," *Trends Food Sci. Technol.*, vol. 42, no. 1, pp. 27–43, Mar. 2015.

[64] A. M. Vivar-Quintana, C. Santos-Buelga, and J. C. Rivas-Gonzalo, "Anthocyanin-derived pigments and colour of red wines," *Analytica Chim. Acta*, vol. 458, no. 1, pp. 147–155, Apr. 2002.

[65] R. B. Boulton, V. L. Singleton, L. F. Bisson, and R. E. Kunkee," *Principles and Practices of Winemaking*. Cham, Switzerland: Springer, 2013.

[66] H.-F. Liu, B.-H. Wu, P.-G. Fan, S.-H. Li, and L.-S. Li, "Sugar and acid concentrations in 98 grape cultivars analyzed by principal component analysis," *J. Sci. Food Agricult.*, vol. 86, no. 10, pp. 1526–1536, 2006.

[67] S. Kumar, K. Agrawal, and N. Mandan, "Red wine quality prediction using machine learning techniques," in *Proc. Int. Conf. Comput. Commun. Informat. (ICCCI)*, Jan. 2020, pp. 1–6.

[68] E. Wiki, "Spearman's rank correlation coefficient," Tech. Rep., 1988.

[69] Y. Zhou and S. Li, "BP neural network modeling with sensitivity analysis on monotonicity based Spearman coefficient," *Chemometric Intell. Lab. Syst.*, vol. 200, May 2020, Art. no. 103977.

[70] B. C. Kuo and F. Golnaraghi, "Design of control system," *Autom. Control Syst.*, 1995.

[71] G. C. Karmakar, R. Das, and J. Kamruzzaman, "IoT sensor numerical data trust model using temporal correlation," *IEEE Internet Things J.*, vol. 7, no. 4, pp. 2573–2581, Apr. 2020.

[72] A. Asuncion and D. Newman. (2007). *UCI Machine Learning Repository*. [Online]. Available: https://archive.ics.uci.edu/ml/machine-learning-databases/wine-quality/

[73] Z. Wu, M. Zhu, Y. Kang, E. L.-H. Leung, T. Lei, C. Shen, D. Jiang, Z. Wang, D. Cao, and T. Hou, "Do we need different machine learning algorithms for QSAR modeling? A comprehensive assessment of 16 machine learning algorithms on 14 QSAR data sets," *Briefings Bioinf.*, vol. 22, no. 4, Jul. 2021, Art. no. bbaa321.

[74] G. Liu, D. Zhou, H. Xu, and C. Mei, "Model optimization of SVM for a fermentation soft sensor," *Expert Syst. Appl.*, vol. 37, no. 4, pp. 2708–2713, Apr. 2010.

[75] S. K. Sunter and N. Nagi, "A simplified polynomial-fitting algorithm for DAC and ADC BIST," in *Proc. Int. Test Conf.*, Nov. 1997, pp. 389–395.

[76] I. Abubakar, S. Khalid, M. Mustafa, H. Shareef, and M. Mustapha, "Calibration of zmpt101b voltage sensor module using polynomial regression for accurate load monitoring," *J. Eng. Appl. Sci.*, vol. 12, no. 4, pp. 1077–1079, 2017.

[77] D. Longstreet. (2015). *Polynomial Models Explained (for the Non-Data Scientist)*. Accessed: Mar. 31, 2023. [Online]. Available: https://www.linkedin.com/pulse/polynomial-models-explained-non-data-scientist-david-longstreet/

[78] H. Akoglu, "User's guide to correlation coefficients," *Turkish J. Emergency Med.*, vol. 18, no. 3, pp. 91–93, Sep. 2018.

[79] S. Yahya and B. Kingsman, ''Vendor rating for an entrepreneur development programme: A case study using the analytic hierarchy process method,'' *J. Oper. Res. Soc.*, vol. 50, no. 9, pp. 916–930, Sep. 1999.
[80] J. Gogodze, ''Innovative advantages ranking: A new approach,'' *Oper. Res. Decisions*, vol. 29, no. 1, pp. 5–15, 2019.
[81] Q. Tao, F. Liu, Y. Li, and D. Sidorov, ''Air pollution forecasting using a deep learning model based on 1D convnets and bidirectional GRU,'' *IEEE Access*, vol. 7, pp. 76690–76698, 2019.

SACHIN K. SEN (Member, IEEE) received the B.Sc. degree in electrical and electronics engineering from the Khulna University of Engineering and Technology, Bangladesh, and the M.Eng. degree in electrical and electronics from The University of Auckland, New Zealand. He is currently a Lecturer of network and cybersecurity with the School of Computing, Electrical and Applied Technology, Unitec Institute of Technology, New Zealand. He is also a Cisco CCIE and has more than 14 years of industry experience in managing enterprise networks and security in New Zealand. His research interests include the cybersecurity of IIoT-enabled infrastructures and applications, cyber-physical systems, data analytics, and mathematical modeling.

GOUR C. KARMAKAR (Member, IEEE) received the B.Sc. degree from the Department of Computer Science and Engineering (CSE), Bangladesh University of Engineering and Technology (BUET), in 1993, and the master's and Ph.D. degrees in information technology from the Faculty of Information Technology, Monash University, in 1999 and 2003, respectively. He is currently an Associate Professor with Federation University Australia. He has published more than 160 peer-reviewed research publications, including 40 international peer-reviewed reputed journal articles and was awarded six best papers in reputed international conferences. He received the prestigious ARC linkage grant in 2011. He has successfully supervised 16 Ph.D. students. His research interests include multimedia signal processing, big data analytics, the Internet of Things, and cybersecurity including trustworthiness measures.

SHAONING PANG (Senior Member, IEEE) received the B.Sc. degree in physics and the M.Sc. degree in electrical engineering from Xinjiang University, China, in 1994 and 1997, respectively, and the Ph.D. degree in computer science from Shanghai Jiao Tong University, Shanghai, China, in 2000. From 2001 to 2003, he was a Research Associate with the Pohang University of Science and Technology (POSTECH), Pohang, South Korea. He is currently an Associate Professor of cybersecurity with the Institute of Innovation, Science and Sustainability, Federation University Australia. Before joining Federation University, he was a Professor of data analytics and the Director of the Center Computational Intelligence for Cybersecurity, Unitec Institute of Technology, New Zealand. His research interests include cybersecurity, data analytics, cloud data and service resiliency, and urban environmental big data.