# Predictive Method Proposal for a Manufacturing System with Industry 4.0 Technologies

Santiago Aguirre[1](✉), Lina Zuñiga[1], and Michael Arias[2]

[1] Pontificia Universidad Javeriana, Bogotá, Colombia
{saguirre,lmzunigaq}@javeriana.edu.co
[2] Universidad de Costa Rica, San Pedro, Costa Rica
michael.arias_c@ucr.ac.cr

**Abstract.** Cyber-physical manufacturing systems with industry 4.0 technologies have the ability to generate real-time data on the behavior of the system in each of its components, so predictions can be generated from this data. This article presents a method for the development of a predictive model where process mining techniques and data mining algorithms are combined. Through the discovery techniques of process mining, a descriptive analysis of the system is carried out to subsequently develop a predictive model with predictive data mining algorithms that provide information on the time remaining for a product that is in process to be completed. This prediction allows decision makers to reconfigure the manufacturing system variables and its schedule to optimize its performance. The method was applied in a production system that is currently installed in the Computer Integration Manufacturing Lab at Pontificia Universidad Javeriana.

**Keywords:** Process mining · Predictive monitoring · Industry 4.0 · Data mining

## 1 Introduction

The Fourth Industrial Revolution, also known as Industry 4.0, is a concept that seeks to increase the efficiency of production systems by making them more flexible [13] through the incorporation of relevant technologies [20]. A comprehensive definition of this concept [6] makes it possible to recognize four fundamental elements and technologies in the industry: smart factories, cyber-physical systems and internet of things (IoT).

Cyber-physical systems are integrations of computing and physical processes that, through integrated computers and networks, monitor and control processes in real time [9]. The abstractions and models that derive from this integration are often used to monitor and control the performance of these systems, in order to subsequently apply tools, techniques and methodologies that allow them to be optimized.

Among the main characteristics of cyber-physical manufacturing systems is an inherent component of dynamism, which makes the diagnosis of the efficiency of the system a real challenge. However, recent results and developments at the technological level have allowed greater availability and affordability of sensors, data acquisition systems

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2022
J. C. Figueroa-García et al. (Eds.): WEA 2022, CCIS 1685, pp. 109–121, 2022.
https://doi.org/10.1007/978-3-031-20611-5_10

and computer networks, which allow the continuous generation of data. In such an environment, cyber-physical systems can be developed and improved by managing data and leveraging system interconnectivity to achieve intelligent, resilient, and self-adaptive machines [10].

The information that is registered in the monitoring and control nodes of the cyber-physical manufacturing systems can be exploited through process mining, which is made up of techniques and algorithms that allow discovering the real execution model of the processes, monitoring key indicators, analyzing them, and looking for ways to improve them [1]. Process mining techniques have been widely applied for the diagnosis, descriptive analysis, and improvement of all types of processes, including manufacturing systems [5, 8, 12]. However, in the literature review, it was found that limited applications of process mining for the development of predictive models in cyber-physical manufacturing systems [4, 11].

The contribution of this article is based on the combination of process mining techniques, which allow the descriptive analysis of a manufacturing system, with data mining techniques and algorithms such as decision trees, that allow developing predictions of the remaining cycle time of a production order to make decisions that contribute to cycle time and resource use optimization. The proposed method was applied at the Computer Integration Manufacturing Lab located at Pontificia Universidad Javeriana, where there is a cyber-physical manufacturing system with industry 4.0 technologies.

This document is organized as follows: Section 2 describes the background and related work. Section 3 describes the manufacturing system on which the predictive method was developed, and Sects. 4, 5, and 6 describe the developed method and its application. The results are presented in Sect. 7, which ends with the conclusions and future work.

## 2 Related Work

In recent years, some studies have been conducted for the application of process mining in manufacturing systems that have the capacity to generate data in real time [5, 8, 12]. These case studies enable the identification of opportunities for improvement in various scenarios involving the transformation of inputs into finished products via a manufacturing system. The most relevant case studies are described below.

In the article by Jiménez et al. [8], the simulation of a flexible manufacturing system installed in the laboratory of the Université de Valenciennes in France is described. This system is based on two machine selection rules under normal and non-normal production conditions subject to disturbances. Through process mining, a descriptive analysis of the system was carried out and it was concluded that these techniques can be used to diagnose the behavior of manufacturing systems and to compare the system performance based on different process configurations and product routing.

On the other hand, Schuh et al. [15] proposed a data-based methodology for process performance analysis in the manufacturing industry. The methodology consists of three steps: (1) extracting performance-related event logs; (2) merging and preparing event logs from multiple sources; and (3) process discovery for performance description and analysis. Its main contribution was the incorporation of process mining into all value chain process networks, including sales, manufacturing, and order fulfillment. An application of the methodology was presented in a real case study corresponding to a small metal-mechanic products company, where the real execution model of the fulfillment process was described, starting from the customer need until product delivery.

Process mining is not only useful for process discovery but can also be used to predict process behavior [1]. A relevant case of this application of the discipline is in supply chain analysis, where a scenario consisting of several independent factories was simulated [16]. In this case, customers requested certain products and each factory had until the end of the day to produce those items. Process mining process discovery algorithms, along with decision analysis techniques, were applied to analyze the decisions each factory made and thus predict what decisions it might make in the future.

The remaining production time in a manufacturing process could be predicted through process mining. Choueiri [4] proposed a hybrid predictive model, which starts from the discovery of the process through a process mining algorithm, from which transition systems and statistical regression models were applied to predict the remaining production times. The model was tested on an artificially created log emulating an industrial environment and on a real manufacturing log. The results showed that the approach provided better precision measures compared to the method applied in a previous development by Van der Aalst et al. [2].

Finally, another example of process mining applications for the prediction of their behavior was developed by Lovera et al. [11], who designed a methodology for the introduction of a predictive model in a manufacturing system, allowing the system to make better decisions. The predictive process mining algorithms available in the Apromore software were used as the basis for the development of the proposed methodology and its implementation in the simulated system.

The main limitations of previous predictive developments are based on their applications that are limited to manufacturing environments with a very limited number of configurations regarding the production process (routes) and the number of different products to be produced. The predictive method proposal developed in this article is aimed at a fully automated cyber-physical flexible manufacturing system, where different options for production routes and product configurations are evaluated.

## 3 Manufacturing System Description

The cyber-physical manufacturing system on which the predictive method was developed is a flexible manufacturing system located at the Engineering Laboratories Building of Pontificia Universidad Javeriana in Bogotá, Colombia. This technology center based on Industry 4.0 technologies seeks the complete digitization of a company's value chains through the integration of data processing technologies, artificial intelligence and IoT sensors. The laboratory has several workstations that include a raw material and finished product warehouse, a conveyor belt for material flow, manipulator robots, and a quality control station by artificial vision.

The manufacturing cell in which the predictive model was applied is made up of five workstations, four made up of a single machine (M2, M3, M4 and M5) and another made up of a storage module (M1). Each module and workstations are configured to perform different manufacturing operations. The workstations are connected through a one-way transport band system. For product movement and assembly, a transport automated guided robot (AGV) is used. This self-propelled resource transports products from station to station and, thanks to a basic behavior system, avoids colliding with other workstations, detects a transfer node (where the production route to follow is decided), and automatically manages speed and stops in front of the workstations.

Sample mobile phones are made in the manufacturing cell, but since the purpose of this manufacturing system is to learn about topics related to industry 4.0, mobile phones are not real. There are 5 types of mobile phones models (A, B, C, D, and E), each one contains a specific type and number of components which are assembled through a sequence of production operations (production routes). Figure 1 shows a representation of the manufacturing system.

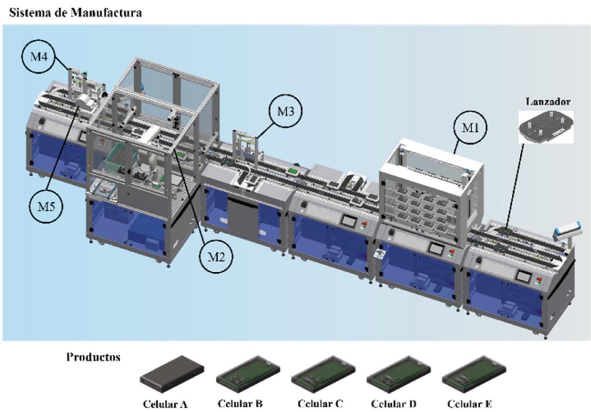


**Fig. 1.** Flexible manufacturing system

The manufacturing system has four interconnected components that allow modeling and controlling the operations in the system: 1) A manufacturing execution system (MES) allows for the definition and execution of production orders, 2) A database comprised of 65 relational tables that allows for real-time access to information on production orders, products, and system configuration, 3) Physical plant: the physical installation of the manufacturing facility, as well as a 4) Simulator program that allows 3D simulation of the manufacturing system.

## 4 Method Description

The main objective of this work is the development of a predictive model in the manufacturing system. To carry it out, a work method was designed and implemented based on the steps of the CRISP-DM method (Cross-Industry Standard Process for Data Mining) that has been tested in different industries to guide data mining related projects [7]. Based on the information in this guide, the following work stages were defined and developed: 1) understanding of the manufacturing system, 2) data understanding, 3) data extraction and preparation, 4) descriptive and predictive model development, 5) model evaluation and 6) monitoring and control.

The development of the first stage consisted of making a characterization of the manufacturing system that would allow a contextualization and understanding of how these components work and are integrated. Having this clear, the definition of manufacturing orders was carried out to cover relevant products and production routes. The quantity and product types in each production order were determined in order to obtain execution information with all possible variations. Subsequently, the production orders were simulated to obtain the metadata and relational tables of the execution of each of the events. In total, 36 production orders were executed, resulting in 710 mobile equipment production cases. In stages 2 and 3, an analysis of the data structure was made, it was determined which ones were going to be used and finally the appropriate treatment was made to obtain the final event log to be used in the next stage. An extract from the final database event log can be seen in Table 1.

**Table 1.** Event log sample.

| Case ID | Activity | Resource | Timestamp Start | Timestamp End | Product | OPos |
|---|---|---|---|---|---|---|
| 206916001 | release a defined part on stopper 1 | high_bay_storage | 2021/05/29 10:04:01 | 2021/05/29 10:04:05 | Celular Sin PCB | 1 |
| 206916001 | feed back cover from magazine | magazine_application_module | 2021/05/29 10:04:50 | 2021/05/29 10:04:51 | Celular Sin PCB | 1 |
| 206916001 | pressing with force regulation | muscle_press_application_module | 2021/05/29 10:04:58 | 2021/05/29 10:05:00 | Celular Sin PCB | 1 |
| 206916001 | print label | labeling_application_module | 2021/05/29 10:05:04 | 2021/05/29 10:05:07 | Celular Sin PCB | 1 |
| 206916001 | store a part from stopper 1 | high_bay_storage | 2021/05/29 10:05:23 | 2021/05/29 10:05:27 | Celular Sin PCB | 1 |
| 206926001 | release a defined part on stopper 1 | high_bay_storage | 2021/05/29 10:04:11 | 2021/05/29 10:04:16 | Celular Sin PCB | 2 |
| 206926001 | feed back cover from magazine | magazine_application_module | 2021/05/29 10:04:53 | 2021/05/29 10:04:54 | Celular Sin PCB | 2 |
| 206926001 | pressing with force regulation | muscle_press_application_module | 2021/05/29 10:05:02 | 2021/05/29 10:05:03 | Celular Sin PCB | 2 |
| 206926001 | print label | labeling_application_module | 2021/05/29 10:05:09 | 2021/05/29 10:05:11 | Celular Sin PCB | 2 |
| 206926001 | store a part from stopper 1 | high_bay_storage | 2021/05/29 10:05:30 | 2021/05/29 10:05:34 | Celular Sin PCB | 2 |

Stage 4 was the most important stage of the method, since, with the application of process mining techniques, a descriptive analysis and a predictive analysis of the system were obtained. The descriptive analysis sought a deeper characterization than the one carried out in stage 1, since relevant data such as performance metrics of the execution of the processes were obtained and analyzed. For its part, predictive analysis seeks, supported by machine learning techniques, to learn from the behavior of the data and build models that make predictions of the remaining process time to finish manufacturing a product.

The predictive models obtained were evaluated in stage 5, based on performance measures of the predictions when making a comparison with data from the real execution based on the test data set. When making the comparison between several models, the one with the best performance metrics was chosen. Finally, in stage 6, a predictive monitoring control panel was developed that interactively allows those in charge of the process to make use of the predictive model in real time and know the status and predictions of the manufacturing cell. In the following two sections, the modeling stage is explained in more detail.

## 5 Descriptive Analysis Through Process Mining

The event log obtained at the end of the development of stage 3 of the method were analyzed through process mining. Celonis software was used as a support tool, which allows integrating process mining with machine learning components for process monitoring [19].

### 5.1 Process Overview

The first part of the descriptive analysis was the analysis of the process cycle time. The review showed that the average manufacturing time for a cell phone is 201 s and that the product that takes the least time to make is cell phone A (it takes around 180 s). Additionally, it was established that the bottleneck is found when the process begins. This is because all launchers stop and cannot continue on the conveyor until the robotic arm of the storage machine picks up or drops off the products as scheduled in the production orders. The next machine that generates an additional delay in the production time of each cell is the robotic assembly cell. It was observed that when the production orders have mobile equipment with plates in consecutive positions, the process time increases considerably.

### 5.2 Process Discovery

After the general review of the process, we proceeded with the exploration of it in order to analyze and understand it. The first thing that was observed when analyzing the activities and the process sequence in each of the cases is that, due to the configuration of the manufacturing cell, the cases always follow an established sequence according to the initial definition of the process. That is, there are no deviations where one passes from one activity to another without following the established normal production route, as can be seen in the discovery diagram of process mining in Fig. 2.

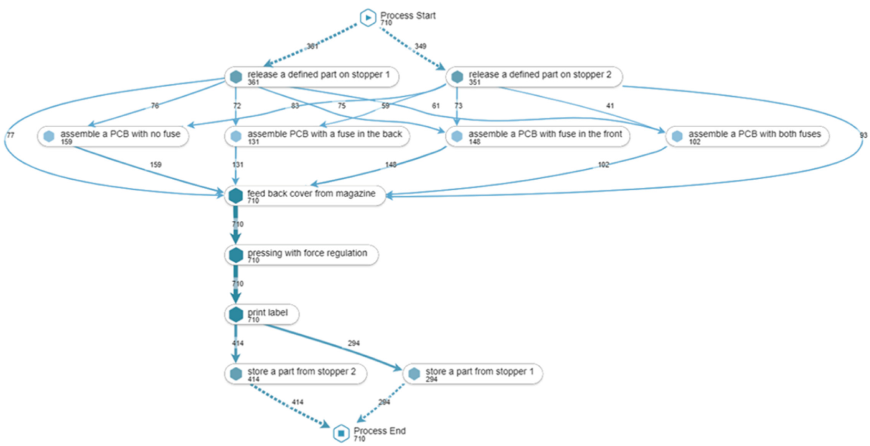


**Fig. 2.** Real process model

With respect to resource usage and the average processing time, it was found that the robotic assembly workstation (M2) is the one that takes the longest with 26.65 s (see Table 2).

**Table 2.** Average machine operation time

| Work center | Average processing time (seconds) |
|---|---|
| High bay storage (M1) | 5.43 |
| Robot assembly station (M2) | 26.62 |
| Magazine application module (M3) | 1.15 |
| Muscle press module (M4) | 1.07 |
| Labeling module (M5) | 2.47 |

### 5.3 Variant Comparison

Given the characteristics of the storage module (M1), it is possible to determine in the production order the starting or ending place of each manufactured product, either on rail 1 (R1) or on rail 2 (R2). This variation in the production orders gives flexibility to the manufacturing cell, since depending on where the product starts and ends, the final time of the production order will vary. Each of the four possible variations was analyzed according to average cycle time, standard deviation, longest cycle time, and shortest cycle time. The result of this comparison (see Table 3) determines that the variation in the process that has better execution times is that of the products that start on rail 2 and end up on rail 1, followed by those that leave and end up on rail 2.

**Table 3.** Product routes variants comparison

| Variant | Average cycle time (seconds) | Standard deviation | Longest cycle time (seconds) | Shorter cycle time (seconds) |
|---|---|---|---|---|
| R1-R1 | 347.99 | 202.29 | 882 | 53 |
| R1-R2 | 293.71 | 180.79 | 889 | 63 |
| R2-R1 | 199.99 | 136.35 | 668 | 45 |
| R2-R1 | 201.22 | 124.7 | 805 | 53 |

## 6 Predictive Analysis

The process flow for the evaluation of the different predictive model based on the data obtained from the 710 cases is shown in Fig. 3. A combination of different modeling techniques were evaluated, which generated several predictive models. Each of them were analyzed and selected based on performance evaluation metrics.

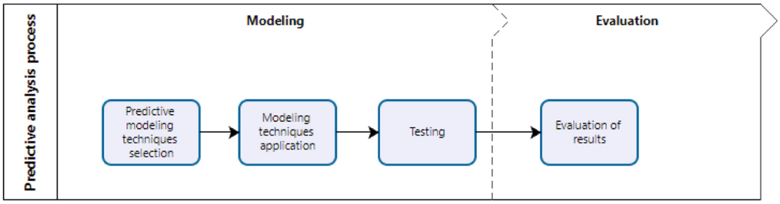


**Fig. 3.** Predictive analysis process flow

## 6.1 Predictive Modeling Techniques

To generate the possible models, a combination of three features of the predictive modeling techniques was used: bucketing method, encoding method and learning method.

- Bucketing. The sequences of the different events in the database are divided into several groups and different models are trained for each of these groups [17].
- Encoding: To train a model, all cases and records that are in the same group must be represented as feature vectors of fixed and unified length [17].
- Learning method: Supervised learning methods were used for model training, in this case: decision trees, random forest and gradient boosting. These methods have been using for predictive process monitoring in previous work [2, 11, 17] showing good results in predicting a variable based on a process event log.

The combination and selection of each of these modeling techniques yielded 36 models. Figure 4 shows these combinations schematically.

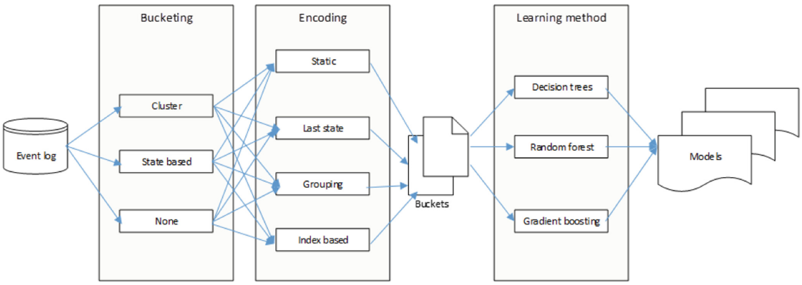


**Fig. 4.** Predictive model generation

## 6.2 Model Comparison and Selection

Python development integrated with Nirdizati (2020) was used for the evaluation of the different models that result from the combination of the predictive method, bucketing and encoding options. Phyton was used because of its features not only when training the data but also because of the future possibility of connecting to the manufacturing cell database and making predictions in real time.

To determine the accuracy of the prediction model, a comparison is made between the predicted value and the real value of each model using the test data set (20% of the initial base). Three performance metrics were used: mean absolute error (MAE), root mean square error (RMSE) and coefficient of determination (R2). Table 4 shows the value of these metrics for each model.

**Table 4.** Prediction model evaluation

| Model | Predictive method | Bucketing | Encoding | MAE | RMSE | R2 |
|---|---|---|---|---|---|---|
| 1 | Decision trees | None | Static | 26.49 | 50.64 | 0.81 |
| 2 | Decision trees | None | Last state | 19.19 | 28.60 | 0.67 |
| 3 | Decision trees | None | Grouping | 25.01 | 53.47 | 0.80 |
| 4 | Decision trees | None | Index based | 18.30 | 38.5 | 0.82 |
| 5 | Decision trees | Cluster | Static | 27.66 | 55.88 | 0.83 |
| 6 | Decision trees | Cluster | Last state | 18.92 | 28.26 | 0.67 |
| 7 | Decision trees | Cluster | Grouping | 24.45 | 52.14 | 0.80 |
| 8 | Decision trees | Cluster | Index based | 18.79 | 39.57 | 0.82 |
| 9 | Decision trees | Last state | Static | 15.44 | 23.09 | 0.52 |
| 10 | Decision trees | Last state | Last state | 15.44 | 23.10 | 0.53 |
| 11 | Decision trees | Last state | Grouping | 15.44 | 23.11 | 0.54 |
| 12 | Decision trees | Last state | Index based | 15.44 | 23.12 | 0.55 |
| 13 | Random forest | None | Static | 53.59 | 121.14 | 0.75 |
| 14 | Random forest | None | Last state | 13.26 | 20.31 | 0.45 |
| 15 | Random forest | None | Grouping | 54.89 | 121.41 | 0.76 |
| 16 | Random forest | None | Index based | 55.32 | 121.57 | 0.76 |
| 17 | Random forest | Cluster | Static | 53.57 | 121.26 | 0.75 |
| 18 | Random forest | Cluster | Last state | 13.26 | 20.31 | 0.45 |
| 19 | Random forest | Cluster | Grouping | 54.89 | 121.41 | 0.76 |
| 20 | Random forest | Cluster | Index based | 55.32 | 121.57 | 0.76 |
| 21 | Random forest | Last state | Static | 15.70 | 23.00 | 0.52 |
| 22 | Random forest | Last state | Last state | 15.48 | 23.14 | 0.53 |
| 23 | Random forest | Last state | Grouping | 15.49 | 23.15 | 0.54 |
| 24 | Random forest | Last state | Index based | 15.50 | 23.16 | 0.55 |
| 25 | Gradient boosting | None | Static | 16.45 | 19.96 | 0.84 |
| 26 | Gradient boosting | None | Last state | 23.59 | 31.34 | 0.00 |
| 27 | Gradient boosting | None | Grouping | 28.66 | 38.18 | 0.84 |
| 28 | Gradient boosting | None | Index based | 8.64 | 12.37 | 0.88 |
| 29 | Gradient boosting | Cluster | Static | 15.16 | 20.02 | 0.75 |
| 30 | Gradient boosting | Cluster | Last state | 15.35 | 20.28 | 0.76 |
| 31 | Gradient boosting | Cluster | Grouping | 42.27 | 68.89 | 0.83 |
| 32 | Gradient boosting | Cluster | Index based | 9.01 | 14.02 | 0.87 |
| 33 | Gradient boosting | Last state | Static | 15.41 | 23.05 | 0.52 |
| 34 | Gradient boosting | Last state | Last state | 15.42 | 23.06 | 0.53 |
| 35 | Gradient boosting | Last state | Grouping | 15.43 | 23.07 | 0.54 |
| 36 | Gradient boosting | Last state | Index based | 15.44 | 23.08 | 0.55 |

To facilitate visualization and comparison of performance measures, a scatter plot graph was constructed (see Fig. 5) in which the value of the root mean square error (RMSE) versus the coefficient of determination (R2) for each of the 36 models can be visualized. The labels of each observation on the graph correspond to the identification number assigned to each model (see Table 4). The coefficient of determination was plotted because, unlike the other two performance measures, it reflects how well the model predicts an outcome. Regarding the other two metrics, although both have the analytical advantage of expressing the error in units of time (the unit of measurement of the response variable), the root mean square error (RMSE) was selected because, unlike the mean absolute error (MAE), in its magnitude a relatively high weight is given to large errors, which could be considered significant in the context of the prediction model.

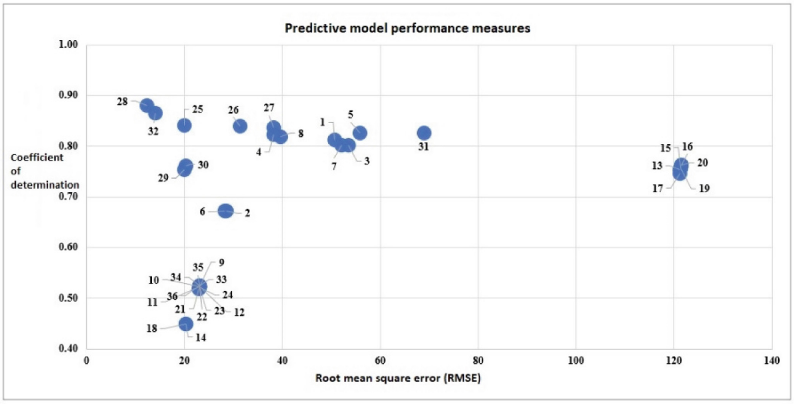


**Fig. 5.** Comparison of predictive models performance measures

When making an overall comparison of the performance metrics between the 36 models, the models identified with labels 28 and 32 are the ones that present the best results. These two models are characterized by the fact that the learning method used was gradient boosting and the coding of the groups was index-based. Given that between these two models, the one with the best performance metrics is the one with the label 28, this is the one selected to be implemented in the manufacturing cell to make predictions in real time.

When analyzing the graph of the remaining process time predicted value and the real remaining process time value of the results produced by the selected model (see Fig. 6), it can be seen that although there are some outliers, most of the data is aligned into the middle of the graph. This leads to the conclusion that most of the predicted values are very close to the real values.

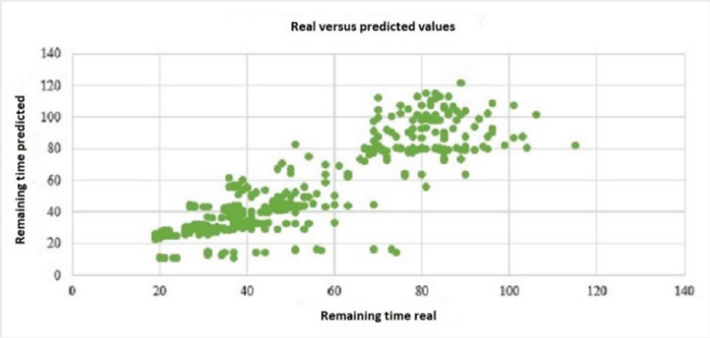


**Fig. 6.** Comparison of real values vs prediction

## 7 Results and Discussion

For the development of the methodology, the production of 710 products in the manufacturing cell was simulated. Based on the event log obtained, it was possible to perform a descriptive analysis using process mining tools that allowed mainly to characterize the behavior of the manufacturing process and to determine the resource restrictions. The storage and the robotic assembly workstations are the resources that generate a considerable increase in process times. In the case of the storage workstation, this occurs because all the products must begin and end there, and sometimes there are wait times due to the limited capacity of the robotic arm that moves the product to the conveyor belt.

The predictive model enables determining the best predictive method based on the combination of bucketing and encoding options to predict the remaining process time. Gradient boosting trees predictive methods had the best performance based on root mean square error (RMSE) and the coefficient of determination (R2).

By combining the results obtained by both analyses, it is possible to improve the way in which production orders are scheduled and make timely decisions that improve the efficiency of the system since, on the one hand, the limitations of the system's resources are known, and on the other hand, through predictive monitoring, it will be possible to determine the estimated time of completion of a product or production order. The foregoing will also allow, as long as there is not a high volume of work in process orders, to determine which should be the next one to be produced and in the same sense, to fulfill established commitments or target times.

## 8 Conclusions and Future Work

Manufacturing environments are being transformed thanks to the fourth industrial revolution called Industry 4.0. Given the flexibility requirements that characterize cyber-physical manufacturing systems, it is essential to develop data analytics methods that leverage the greater availability of data generation and acquisition in order to analyze and improve key performance manufacturing measures. Based on this, through the proposed method developed in this work, a predictive model was obtained for a computer integrated manufacturing system that allows the prediction of the remaining processing time to finish a product that is in process to improve the decision-making.

Although the scope of the work was limited to the development of a single predictive model, through the implementation of the developed method, it is possible to train models to predict other important system metrics such as the production capacity indicators. By predicting the number of mobile equipment that could be manufactured per hour given certain conditions, it would be possible to determine the efficiency of the system and would allow the process leader to make better decisions related to the production order sequencing for optimizing processing times. For future work, it is proposed to use the results of the prediction of the remaining time of the production orders with optimization models for the sequencing of the production orders that allow changes to be made in real time and thus optimize the times and use of resources.

Finally, another promising research perspective is related to the continuous incorporation of real-time data for enhancing the prediction model. The model training could be automated with the variation of certain parameters and its subsequent evaluation through performance metrics. This would allow having better predictors each time, since the predictive model would have more data to learn from.

## References


1. Aalst, W.V.D.: Process Mining - Data Science in Action. Springer, Heidelberg (2016). https://doi.org/10.1007/978-3-662-49851-4
2. Aalst, W.V.D., Schonenberg, M.H., Song, M.: Time prediction based on process mining. Inf. Syst. **36**(2), 450–475 (2011)
3. van der Aalst, W., et al.: Process mining manifesto. In: Daniel, F., Barkaoui, K., Dustdar, S. (eds.) BPM 2011. LNBIP, vol. 99, pp. 169–194. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-28108-2_19
4. Choueiri, A.C., Sato, D.M.V., Scalabrin, E.E., Santos, E.A.P.: An extended model for remaining time prediction in manufacturing systems using process mining. J. Manuf. Syst. **56**, 188–201 (2020)
5. Dreher, S., Reimann, P., Gröger, C.: Application fields and research gaps of process mining in manufacturing companies. INFORMATIK **2020** (2021)
6. Hermann, M., Pentek, T., Otto, B.: Design principles for industries 4.0 scenarios. In: 2016 49th Hawaii international conference on system sciences (HICSS), pp. 3928–3937. IEEE (2016)
7. IBM: CRISP-DM help overview (2020). https://www.ibm.com/docs/en/spssmodeler/SaaS?topic=dm-crisp-help-overview. Accessed 2 May 2022
8. Jimenez, J.F., Zambrano-Rey, G., Aguirre, S., Trentesaux, D.: Using process mining for understating the emergence of self-organizing manufacturing systems. IFAC-PapersOnLine **51**(11), 1618–1623 (2018)
9. Lee, E.A.: Cyber physical systems: design challenges. In: 2008 11th IEEE International Symposium on Object and Component-Oriented Real-Time Distributed Computing (ISORC), pp. 363–369. IEEE (2008)
10. Lee, J., Bagheri, B., Kao, H.A.: A cyber-physical systems architecture for industry 4.0-based manufacturing systems. Manuf. Lett. **3**, 18–23 (2015)
11. López Castro, L., Martínez, S., Rodriguez, N., Lovera, L., Santiago Aguirre, H., Jimenez, J.-F.: Development of a predictive process monitoring methodology in a self-organized manufacturing system. In: Trentesaux, D., Borangiu, T., Leitão, P., Jimenez, J.-F., Montoya-Torres,

J.R. (eds.) SOHOMA 2021. SCI, vol. 987, pp. 3–16. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-80906-5_1

12. Lorenz, R., Senoner, J., Sihn, W., Netland, T.: Using process mining to improve productivity in make-to-stock manufacturing. Int. J. Prod. Res. **59**(16), 4869–4880 (2021)
13. Lozano, C.V., Vijayan, K.K.: Literature review on cyber physical systems design. Procedia Manuf. **45**, 295–300 (2020)
14. Nirdizati Org.: Why Nirdizati? (2020). http://nirdizati.org/why-nirdizati/. Accessed 30 April 2022
15. Schuh, G., Gützlaff, A., Schmitz, S., van der Aalst, W.M.: Data-based description of process performance in end-to-end order processing. CIRP Ann. **69**(1), 381–384 (2020)
16. Shmueli, G., Bruce, P.C., Yahav, I., Patel, N.R., Lichtendahl, K.C. Jr.: Data Mining for Business Analytics: Concepts, Techniques, and Applications in R. Wiley (2017)
17. Teinemaa, I., Dumas, M., Rosa, M.L., Maggi, F.M.: Outcome-oriented predictive process monitoring: review and benchmark. ACM Trans. Knowl. Discovery Data (TKDD) **13**(2), 1–57 (2019)
18. Van Dongen, B., van Luin, J., Verbeek, E.: Process mining in a multi-agent auctioning system. In: Proceedings of the 4th International Workshop on Modelling of Objects, Components, and Agents, Turku, pp. 145–160 (2006)
19. Veit, F., Geyer-Klingeberg, J., Madrzak, J., Haug, M., Thomson, J.: The proactive insights engine: process mining meets machine learning and artificial intelligence. In: BPM (Demos) (2017)
20. Yilmaz, S.E.: Overcoming the technology myopia of industry 4.0 (2020). https://eds-b-ebscohostcom.ezproxy.javeriana.edu.co/eds/pdfviewer/pdfviewervid=5&sid=9b9034fdd93c-405b-866c-a1fe06e1d94a%40sessionmgr101. Accessed 30 April 2022