Data-driven Discovery of Manufacturing Processes and Performance from Worker Localisation

Ayse Aslan[1[0000−0003−2974−8314]], Hanane El-Raoui[2[0000−0002−9079−3248]], Jack Hanson[3[0000−0001−5105−8497]], Gokula Vasantha[1[0000−0002−5479−6134]], John Quigley[2[0000−0002−7253−8470]], and Jonathan Corney[3[0000−0003−1210−3827]]

[1] Edinburgh Napier University, EH10 5DT, UK
[2] University of Strathclyde, G1 1XQ, UK
[3] The University of Edinburgh, EH8 9YL, UK
a.aslan@napier.ac.uk (Corresponding author)

Abstract. In complex manufacturing industries that are not fully automated and involve human workers it is important to identify deviations from the planned production schedule and locate bottlenecks for improved efficiency. This is not an easy task as it requires data on how workers are actually performing the manufacturing activities. Ultra-wideband (UWB) tags, which are sensors that track movement, can be used to collect this data. Previous research has mostly focused on using these sensors to detect faults and anomalies and to ensure worker safety. However, this paper presents a method for using UWB data to discover process models of manufacturing activities using process mining techniques. We applied our method to a real assembly line with UWB data and found deviations from the prescribed process steps and bottlenecks in the assembly line, which indicated that the first assembly step can take twice as much time compared to other steps.

Keywords: manufacturing process optimisation, industrial productivity, process mining, indoor positioning systems

1 Introduction

In many manufacturing industries where it is not economically feasible to implement automated production lines human workers have to interact with constantly changing environment, consequently deviations from the planned production schedule are common. To improve efficiency, it's important for manufacturers to identify and understand these deviations in order to find bottlenecks in the manufacturing process. However, this can be a challenging task, as it requires real-time data on how workers are actually performing the manufacturing activities. By analysing this data, manufacturers can identify areas where the process is deviating from the plan and take steps to improve efficiency.

Sensors on the shop floors are commonplace in the Industry 4.0 era. They allow tracking of workers, materials, products, and machine states in real-time.

The information they provide gives an opportunity to infer information about a dynamically changing manufacturing state and performance, more accurately than static models or general historical trends (that do not take into account variations in the manufacturing processes). With the help of data science methods, sensor data can be exploited to predict and analyse manufacturing processes within a purely data-driven approach [1].

In this paper, we demonstrate the potential for using localisation sensors that track the positions of workers in a manufacturing environment to discover data-driven process models of manufacturing activities and performance. The derived process models can be used to identify how human workers deviate from prescribed process steps and detect bottleneck manufacturing activities by predicting performance metrics such as mean duration of activities. Our methodology involves extracting event logs from worker localisation data and applying process mining techniques to generate system models that can be used to analyse a range of manufacturing performance measures. To demonstrate the applicability of our approach, we use the real-world dataset provided in [2] that contains the positions of 6 assembly workers in our case study. Our methodology identifies where and how the process deviates from the prescribed assembly steps and their order, and which manufacturing activities are the bottlenecks in the assembly line. Overall, our approach shows the potential for using data-driven process models from worker localisation sensor data to improve the efficiency of non-fully automated manufacturing systems.

The rest of the paper is organised as follows. Section 2 reviews the related literature and highlights how our work differs from the existing studies. In Section 3, we present our data-driven approach, which we apply in a case study in Section 4. Lastly, Section 5 concludes the paper.

2 Literature Review

Performance tracking and identification of deviations and bottlenecks in manufacturing have received significant attention from researchers because of their relevance to improving productivity. Researchers have applied various methods to detect bottlenecks and deviations in manufacturing processes. [3] applied the "Turning Point" method to detect bottlenecks in a serial production line using data on production line blockage, starvation probabilities and buffer content. [4] also used this well-recognised method to non-serial production lines based on time logs of stations which provide data on the status of stations. [5] applied Bayesian inference to analyse the probability of deviating from the scheduled target on production output with various production state data including processing times, instances of production shortage/overage and availability of buffers. Based on the data on machine states collected through a Manufacturing Execution System, [6] proposed a data-drive prognostic algorithm using the active-period bottleneck analysis theory. Machine learning methods such as neural networks are also used for this purpose [7]. All these studies use production state data that can be more directly linked to manufacturing performance, such as the status of work stations, machines, or product flow. In contrast, our paper proposes the use of worker movement data to derive performance measures and identify deviations and bottlenecks in manufacturing processes.

Existing studies using localisation data in manufacturing environments have focused on assessing the accuracy of different localisation technologies such as comparing ultra-wideband (UWB) technology to motion capture technology [2] or ultrasound technology [8], filtering techniques to reduce positioning error in localisation data [9], and monitoring the safety of the work environment with the addition of tracking workers' breathing patterns [10] and with vibration sensors to detect abnormality (e.g. falling of a heavy object) [11]. In this work, similar to [12], we focus on predicting performance of manufacturing activities from localisation data. In [12], authors additionally incorporate the product information in their analysis. Our data-driven approach on the other hand utilises only localisation data on workers' positions, without tracking products. This allows us to discover process models without the need for additional tracking systems.

Process mining techniques have been widely applied in the manufacturing industry for identifying process deviations and detecting bottlenecks, with manufacturing being the third largest application domain for these techniques [13], as demonstrated in previous studies [14,15]. Process mining applications in manufacturing typically rely on event logs of activities performed on specific products, which are often generated using IT tracking systems. [14] uses an order status tracking system to compose event logs on ordered products and the activities performed on them. [15] relies on code readers placed on machines that automatically scan product IDs and record a timestamp when they process them in their event logs. However, in our approach, we extract event logs based on worker localisation data without tracking products.

3 Methodology

Let us consider the manufacturing process of a product that consists of a number of activities $i = 1, 2, .., I$ performed in designated activity zones $k = 1, 2, .., I$ on a shop floor by a number of workers $j = 1, 2, ..., J$. We suppose that through a localisation sensing technology, such as UWB tags, the position of each worker j on the floor is tracked in real-time, showing the position in two dimensions (x, y) according to a chosen reference point in the shop. Using the timestamped worker position data supplied through sensors, we generate models of the manufacturing process and its performance in a purely data-driven approach using process mining techniques. The methodology for this is shown in Figure 1. The steps of our methodology are discussed in the following subsections.

Position Data Synchronisation: The UWB tags take a new measurement every 100 ms. However, the position data points taken from workers can be asynchronous, and during some periods the data might be missing (e.g. signal getting blocked because of the metal objects near the worker). To synchronise the data of all workers, we select seconds as the unit of time period (t) because we do not expect any significant change of worker position on the shop floor in

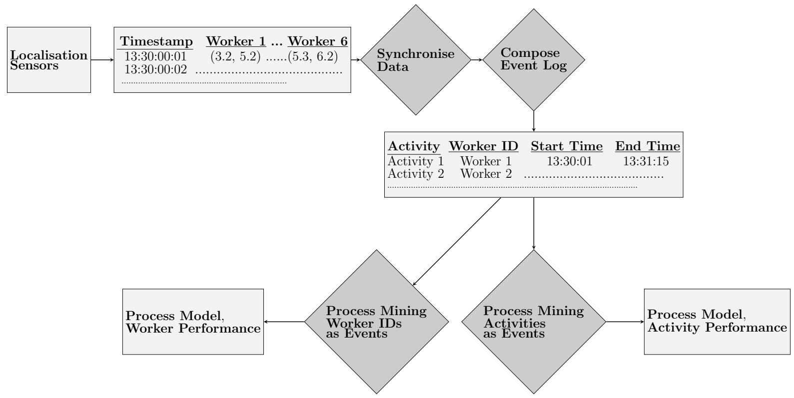

Fig. 1: Methodology: From worker localisation data to manufacturing processes and performance analysis

a second. When there are multiple position data points taken within the same second from a worker, we take the average of those points as the position of the worker in that second. Note that this averaging technique is also a smoothing function over the measurement errors. On the other hand, when there are periods during which no position data points are recorded for a worker, we fill the missing position data with the last known position of the worker.

Composing Event Logs: To apply process mining techniques we need event logs [13] that show what activities are performed by workers and when the activities start and end. We extract an event log from the second-by-second position dataset that we synchronised. In doing this, we assign activities to workers based on their proximity to activity zones, and for how long they were in a particular zone. To consider that a worker is performing a specific activity based on his/her position, we need to take into account whether the time spent in a particular zone is long enough. For example, if the time spent is very short, and the worker is in zone k at second t and in some other zone at second $t+1$, this could be because the worker is walking. More specifically, by using a parameter $\tau > 0$, we consider that worker j is performing activity i at some period t if the position of the worker was in activity j's zone at any period $t^{'} = t, t-1, \ldots, t-\tau$. Event logs in process mining are often accompanied with case IDs that relate to specific jobs on which activities are performed [13]. In this paper, we consider a manufacturing environment in which only workers are tracked, and not the products. Nevertheless, by exploiting the fact that manufacturing tasks are often repetitive in time, namely, by considering that the data generated from a work shift would all relate to a product, which is the workflow of a shift, we can apply process mining techniques.

Process Discovery and Performance Analysis with Process Mining: There are various process model discovery algorithms such as Alpha Miner, Heuristic Miner and Inductive Miner that can be applied to discover a process model from an event log. As in [16], here, we choose to apply the Inductive

Miner, which is an improvement on the Alpha and Heuristic algorithms. We use the implementation in ProM 6.11 (https://www.promtools.org/). To derive a process model of activities we first consider the activities as events. By comparing the model mined from data to the supposed process model, we can identify the deviations. Additionally, to derive insights into the dependencies between workers and their performance, we then let the worker IDs be the activity labels. The mined process models generated by Inductive Miner are annotated with performance information such as waiting and sojourn times which provide insights into bottleneck processes.

4 Case Study

We take the assembly line described in [2] that relates to the manufacturing of tricycles by completing six activities with six workers. In Figure 2, we present the supposed model of this process among activities where arrows between activities represent the precedence relations. The supposed model assigns each activity to a specific worker. From now on, we represent worker j as the worker who is responsible for undertaking activity $i = j$.

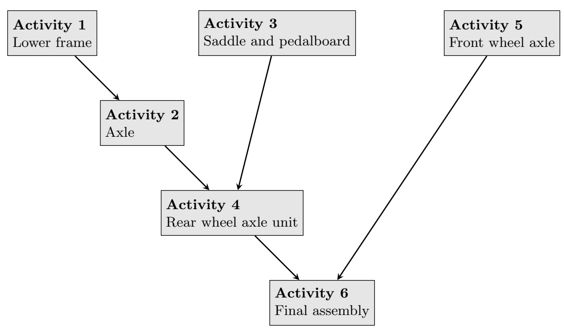

Fig. 2: The assembly line activities for manufacturing tricycles

In this assembly line, the workers' positions on the shop floor are tracked through UWB and motion capture technologies during a three-hours long shift. In our case study, we take into account the UWB data that provides positions in two dimensions because the motion capture system provides fewer data points (because it has been kept running for a shorter duration, around two hours). The coordinates of the UWB positions are measured in meters by considering the anchor position (0, 0) as the reference point in the shop.

We first synchronise this raw UWB data. Figure 3 shows the scatter plot of the position data coming from one of the workers after synchronisation. For each activity, there is a designated rectangular activity zone (this is called a rig in [2]) in the shop. In Figure 4, we show how we divide the shop floor into six activity zones. This is in accordance with the rig locations and shapes as reported in [2].

We then compose an event log based on the positions of workers being inside particular zones for some time by fixing τ to 60 seconds.

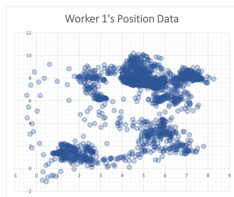

Fig. 3: Worker 1's position data (after synchronisation)

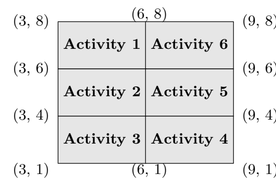

Fig. 4: Locations of activity zones on the shop floor

Before we apply process discovery algorithms, we explore the event log. Figure 5 shows the start times of the activity events performed by six workers. First thing we remark is that there are deviations from the supposed process flow. In particular, we note that worker 2 joins worker 1 for activity 1 during the early stages of the shift. Also, we see that worker 4 sometimes takes on the responsibility of activity 5, and worker 6 helps in activity 2 and 3. The reason for why these workers deviate in this manner could be related to the fact that activities assigned to workers 2, 4 and 6 are dependent on the outputs from earlier assembly stages (see Figure 2), consequently, they cannot be commenced on their own. Moreover, we note that some workers perform activities continuously throughout the shift, worker 5 for instance. However, we see that usually the activities performed by workers are clustered and that there are breaks to activities. The reason for the breaks could be about the completion of the current product and switching to another, as workers complete in total six tricycles during this shift.

4.1 Process Models of Activities and Workers

We apply the Inductive Miner on the event log with activities as event labels to discover the process flow of work among activities in a work shift. The quality of mined process models are often evaluated according to their fitness, precision, simplicity and ability to generalize [13]. These quality dimensions are in line with the objective of finding a balance between overfitting and underfitting that any prediction model strives to achieve. The Inductive Miner can be tuned to allow noise under a certain threshold, and by varying this parameter, models of different fitness and complexity can be discovered from an event log. Specifically, lower the noise threshold, higher the fitness will be. However, high fitness might overfit to the data and can result in complex models that do not reveal important patterns. Under a 30% noise threshold the mined model from our event log has a high fitness of above 91%. However, the resulting model is highly complex, as we present in Figure 6.

In this paper, we instead choose the process model mined under the 40% noise threshold that resulted in much simpler and meaningful representation of

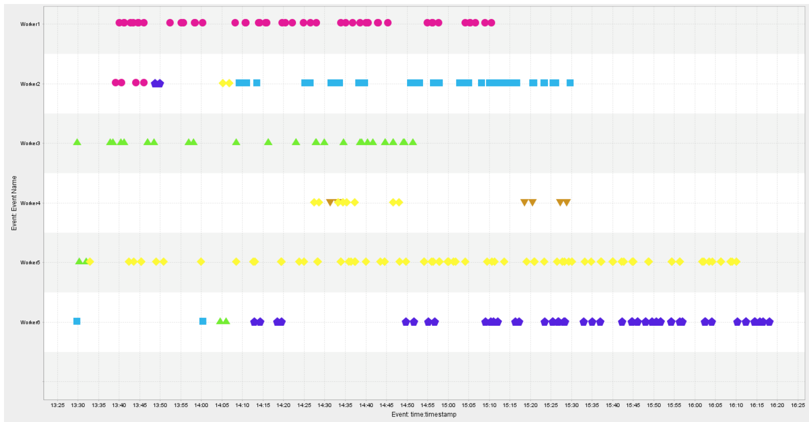

Fig. 5: Timeline of the activities performed by the workers where points mark the start time of the activities (Activity 1: circle, Activity 2: square, Activity 3: up-pointing triangle, Activity 4: down-pointing triangle, Activity 5: diamond, Activity 6: pentagon)

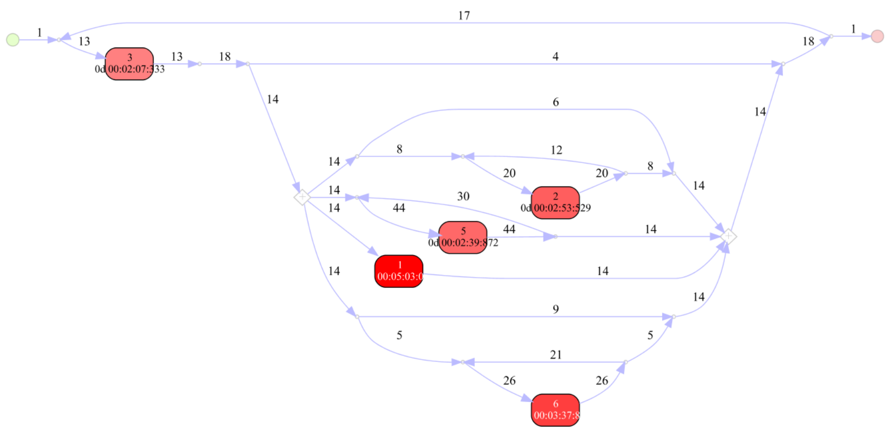

Fig. 6: High-fitness but complex process model of activities with sojourn times

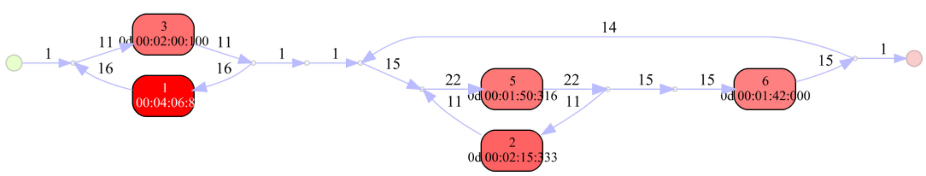

Fig. 7: Process model of activities with sojourn times

the workflow, as illustrated in Figure 7. When the fitness threshold is set high, the derived process model will include a large number of traces observed in the data. This will result in the replication of a variety of patterns in the derived models. When comparing the model in Figure 7 to the high-fitness model in Figure 6, we can see that there are many possible paths among the activities (from the start point to the exit point) in the model shown in Figure 6 that are not possible to observe in the model shown in Figure 7. In both models, any path starts with activity 3, but in the model in Figure 7, the second activity can either be activity 1 or 5, while in the model in Figure 6 it can also be activity 2 or 6.

Numbers above the edges in our process model given in Figure 7 show the frequencies, which are important to understand transitions between activities. Note that not all paths are shown in the model representation. These ones are omitted due to their insignificance. Note also that the model is annotated with average sojourn times spent on activities, where the activities taking the longest to process are coloured redder. According to this model, it seems that activity 1 is the bottleneck of the process, while activities 5 and 6 are performed quicker. Specifically, we see that the sojourn time of activity 1 is more than twice of activities 5 and 6. In terms of compliance, we see that the model discovered from data mostly matches the features and precedence constraints of the supposed process flow given in Figure 2. However, we observe deviations. For example, according to the mined model, activity 5 is performed after activities 1 and 3, while there is no such dependency indicated in the supposed model. Note that the fact that worker 5 is busy with activity 3 during the early stages of the shift (see Figure 5), instead of performing activity 5, could be the reason why we observe this in the mined process model.

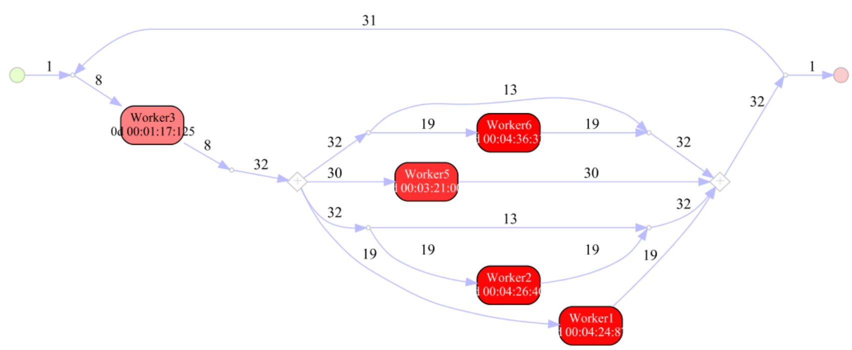

Fig. 8: Process model of workers with sojourn times

In Figure 8, we mine a process model from the event log by considering worker IDs as event labels. We see that there are discrepancies among workers with respect to how long their activities take in the shift. This could be because the workloads of activities are not comparable. Specifically, we observe that workers 1, 2 and 6 work the longest to complete their activities, while activities of worker 3 are much shorter. Note that this is evidenced in the data; in Figure 5 we can see how worker 3 finishes earlier than everyone else.

5 Conclusion and Discussion

This study demonstrates the potential for using localisation sensors that continuously track the positions of workers in a manufacturing environment to discover data-driven process models of manufacturing activities and performance through process mining techniques. Our methodology involves extracting event logs from worker localisation data (e.g. provided by ultra-wideband sensors), applying process mining techniques to the event logs to discover data-driven process models, and deriving manufacturing performance measures. We demonstrate the effectiveness of our methodology by applying it to a real-world assembly line with UWB-based worker localisation data. Our methodology identifies deviations from prescribed assembly steps, as well as bottlenecks in the assembly line. For example, we find that the first assembly step can take more than twice as much time as the last two steps. Overall, our approach shows how data coming from worker localisation sensors can be exploited to discover data-driven process models of manufacturing activities, which then can be used to improve the efficiency of manufacturing systems, which can be particularly valuable for complex systems that are not fully automated. Given that process mining is mainly applied for processes where orders/products are tracked [14,15], the success of our methodology shows that these techniques can also be used when workers' locations are tracked, without tracking orders/products. Our research provided in this paper has some limitations. Our major limitation is on the size of the dataset which we used in our case study. This dataset contains localisation sensor data collected from six assembly workers during a three-hour long shift. However, process models derived from data collected over a longer period may represent the actual manufacturing process more accurately. The extended version of this research will further benefit from the position data to inform and enhance the manufacturing process through re-designing of the layout. Specifically, we will use the position data to identify collaboration relationships between workers and incorporate this to optimise the locations of the activity zones on the shop floor that will minimise the distance travelled by workers.

Acknowledgments This work was supported by the Engineering and Physical Sciences Research Council, UK [grant number EP/V051113/1 -Productivity and Sustainability Management in the Responsive Factory].

References

1. Usuga Cadavid, J.P., Lamouri, S., Grabot, B., Pellerin, R., Fortin, A.: Machine learning applied in production planning and control: A state-of-the-art in the era of Industry 4.0. J. Intell. Manuf. 31, 1531–1558 (2020). https://doi.org/10.1007/s10845-019-01531-7

2. Delamare, M., Duval, F., Boutteau, R.: A new dataset of people flow in an industrial site with UWB and motion capture systems. Sensors 20(16), 4511 (2020). https://doi.org/10.3390/s20164511
3. Li, L., Chang, Q., Ni, J.: Data driven bottleneck detection of manufacturing systems. Int. J. Prod. Res. 47(18), 5019–5036 (2009). https://doi.org/10.1080/00207540701881860
4. Lai, X., Shui, H., Ding, D., Ni. J.: Data-driven dynamic bottleneck detection in complex manufacturing systems. J. Manuf. Syst. 60, 662–675 (2021). https://doi.org/10.1016/j.jmsy.2021.07.016
5. Arashpour, M., Wakefield, R., Blismas, N., Maqsood, T.: Autonomous production tracking for augmenting output in off-site construction. Autom. Constr. 53, 13-21 (2015). https://doi.org/10.1016/j.autcon.2015.03.013
6. Subramaniyan, M., Skoogh, A., Muhammad, A.S., Bokrantz, J., Bekar, E.T.: A prognostic algorithm to prescribe improvement measures on throughput bottlenecks. J. Manuf. Syst. 53, 271-281 (2019). https://doi.org/10.1016/j.jmsy.2019.07.004
7. Huang, H., Wang, W., Ren, S., Zhong, R.Y., Jiang, J.: A proactive task dispatching method based on future bottleneck prediction for the smart factory. Int. J. Comput. Integr. 32(3), 278-293 (2019). https://doi.org/10.1080/0951192X.2019.1571241
8. Crețu-Sîrcu, A.L., Schiøler, H., Cederholm, J.P., Sîrcu, I., Schjørring, A., Larrad, I.R., Berardinelli, G., Madsen, O.:Evaluation and comparison of ultrasonic and UWB technology for indoor localization in an industrial environment. Sensors 22(8), 2927 (2022). https://doi.org/10.3390/s22082927
9. Barbieri, L., Brambilla, M., Trabattoni, A., Mervic, S., Nicoli, M.: UWB Localization in a smart factory: Augmentation methods and experimental assessment. IEEE Trans. Instrum. Meas. 70, 1–18 (2021). https://doi.org/10.1109/TIM.2021.3074403
10. Islam, F.B., Lee, J-M., Kim, D-S.: Smart factory floor safety monitoring using UWB sensor. IET Sci. Meas. 16(7), 412–425 (2022). https://doi.org/10.1049/smt2.12114
11. Nwakanma, C.I., Islam, F.B., Maharani, M.P., Lee, J-M., Kim, D-S.:Detection and classification of human activity for emergency response in smart factory shop floor. Appl. Sci. 11(8), 3662 (2021). https://doi.org/10.3390/app11083662
12. Ruppert, T., Abonyi, J.: Software sensor for activity-time monitoring and fault detection in production lines. Sensors 18(7), 2346 (2018) https://doi.org/10.3390/s18072346.
13. dos Santos Garcia, C., Meincheim, A., Faria, Junior E.R., Dallagassa, M.R., Sato, D.M.V., Carvalho, D.R., Santos, E.A.P., Scalabrin, E.E.: Process mining techniques and applications – A systematic mapping study. Expert Syst. Appl. 133, 260–295 (2019). https://doi.org/10.1016/j.eswa.2019.05.003
14. Lugaresi, G., Ciappina, A.D., Rossi, M., Matta, A.: Exploiting a combined process mining approach to enhance the discovery and analysis of support processes in manufacturing. Int. J. Comput. Integr., 1–21 (2022). https://doi.org/10.1080/0951192X.2022.2090024
15. 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). https://doi.org/10.1080/00207543.2021.1906460
16. Pradana, M.I.A., Kurniati, A.P., Wisudiawan, G.A.A.: Inductive miner implementation to improve healthcare efficiency on Indonesia national health insurance data. In: 2022 International Conference on Data Science and Its Applications (ICoDSA), pp 239–244 (2022). https://doi.org/10.1109/ICoDSA55874.2022.9862837