# Predicting Quality Issues in Manufactured Goods by means of Process Mining in Enterprise Resource Planning Systems

DARRELL TUFTO, University of Twente, The Netherlands

Process mining techniques have proven effective in analyzing and improving complex business processes, especially in addressing bottlenecks. However, current research predominantly focuses on historical data analysis. This paper proposes an approach that combines process mining with machine learning techniques to explore predictive analysis in real-data environments. The methods use historical data from an Enterprise Resource Planning system to cluster event traces by their process features and conformance to identify patterns associated with high and low-quality manufactured goods.

Additional Key Words and Phrases: Process Mining, Conformance Checking, ERP, Manufacturing Predictions

## 1 INTRODUCTION

The manufacturing industry has undergone a paradigm shift in recent years, increasingly relying on data-driven approaches to enhance operational efficiency and product quality. One of these data-driven solutions is in the form of an Enterprise Resource Planning (ERP) system. These ERP systems enable the organization to integrate all the primary business processes to enhance efficiency [5]. However, little attention is devoted to process monitoring and improvement. This is where process mining comes into play. Process mining offers profound capabilities in analyzing, monitoring, and improving complex business processes. It is the method of distilling a structured process description from a set of real executions [4], and with that, bridges the gap between data-rich systems and actionable insights, making it an invaluable tool in the context of ERP systems used within the manufacturing sector. However, predicting issues still to come in manufactured goods by applying process mining techniques has not been extensively explored.

One of those issues that we aim to answer is a client's complaint about the quality of a particular good; "How can we develop a predictive model to anticipate complaints about manufactured goods?"

Many techniques and models have been created to analyze the features of products, semi-finished products, and the data produced by the machines used to answer that question [20]. However, what if one does not look at the features of the data described above but instead at the overall process?

In non-manufacturing processes, it has been shown that:

- There is a relation between processes not conforming to a model created by process mining algorithms and mistakes being made in such process [9];
- Process features, such as *elapsed time* and *number of executions of activities* in a process, often have a relation with all kinds of data points (e.g., costs incurred) [9].

*TScIT 40, February 2, 2024, Enschede, The Netherlands*


Applying that to manufacturing, we devise the following research questions:

(1) Which process discovery algorithm most effectively identifies non-conformant processes that may indicate quality issues in manufactured goods?
   (a) Which process discovery algorithm has the highest fitness and precision scores in real-life data from an ERP system used by a manufacturing company?
   (b) How accurately are non-conformant processes identified as defects in manufactured goods using conformance-checking techniques when evaluated regarding precision and recall metrics?
(2) How effectively do process features classify manufactured goods with and without complaints, as measured by precision and recall metrics?

By answering these research questions, we can then answer how we should create a predictive model to predict whether a manufactured good will receive complaints ahead of time.

## 2 RELATED WORK

This section discusses work related to our research, highlighting studies on process discovery, conformance checking, and prediction methods within the field of process mining. Additionally, it addresses approaches to reducing complexity in real-life event logs.

*Process Discovery.* Research done by de Weerdt et al. [10] has compared different discovery algorithms under different types of datasets and situations using various metrics to evaluate the produced models. More recently, split-miner [6] has been developed as a process discovery method, which performs well on synthetic datasets. Notably, only producing sound models with high scores for both recall and precision metrics, with low execution times. This research shows that the HeuristicsMiner and ILP-miner have the highest precision and fitness within real-life datasets, which is why they will also be used in this study.

*Conformance Checking.* Bauer et al. [11] did a comprehensive literature review of state-of-the-art conformance-checking techniques. Resulting in an overview of different studies that used different modeling languages, perspectives, algorithm types, and quality metrics. The *Conformance Checker* plugin implemented by Rozalin et al. [15] will be used for this research.

*Prediction within Process Mining.* Process mining techniques are considered backward-looking, but forward-looking methods can be created by extracting features from event logs with recurring problems within processes to develop a predictive model with standard machine-learning techniques ranging from regression and decision trees to neural networks to explain the target feature in terms of descriptive features [1]. In 2015, de Leoni et al. [9] introduced a package called *Perform Predictions of Business Proces Features* into the open-source process mining framework ProM [18] to do such analyses. The package creates a generic framework where any characteristic can be explained in terms of correlations with any set of other characteristics[9]. Resulting in a set framework to create decision trees and linear regression models on process features.

*Reducing Complexity.* A well-known challenge with real-life data is the inherent complexity of the dataset. Bose et al. [8] have described a method for clustering traces within event logs to simplify the complexity of real-life processes, facilitating easier comprehension. Bauer et al. [7] demonstrated a technique for approximating conformance results to streamline conformance checking. This is achieved by determining whether the trace currently undergoing conformance checking is similar to a previously evaluated trace. By employing this method of estimating conformance, it becomes possible to evaluate more traces for conformance while requiring less computational effort.

## 3 METHODOLOGIES

This section outlines the methodologies used to address the research questions. The research comprises an analysis of a synthetic dataset and a case study employing the same methods on a dataset from Lienesch B.V. The study is conducted using the CRoss Industry Standard Process for Data Mining [19] (CRISP-DM), which is a step-by-step process that guides data mining and, by extension, process mining research. This research is confined to the algorithms implemented in ProM [18], an open-source process mining framework.

### 3.1 Business Understanding

For the case study, a comprehensive analysis will be done based on interviews with relevant stakeholders and registered processes within the company. The process, as explained in the interview, will result in a visualization in the form of a Petri net [1]. A Petri net is chosen because the modeling requirements analysis tools for process mining are limited to Petri nets as input [9, 14].

### 3.2 Data Understanding

With an in-depth understanding of the business, we will analyze the event logs table in the company's ERP system to identify activities specific to the production process. Since ERP systems encompass production, sales, human resources, and logistics data, a robust understanding is crucial to filter out irrelevant activities effectively.

### 3.3 Data Preparation

Building on this understanding, an SQL query will be developed to export these logs, including only the relevant activities and features for use within the ProM framework. For the case study, data must be exported from the Microsoft SQL database to a CSV file, which can then be imported into ProM. The synthetic dataset [12] is already supplied in a CSV file.

Once the data is modeled, it may be necessary to apply appropriate pre-processing techniques to ensure better model fitting based on an evaluation of such models' precision and fitness metrics.

### 3.4 Modeling

With the data in the XES format, we can start modeling our data. As mentioned in section 2, much research has been done into the state-of-the-art process discovery methods and conformance-checking techniques. According to the research of de Weerdt et al. [10], the best process discovery algorithms for real-life event logs are HeuristicsMiner and ILP-miner [10], with the former having higher scores in precision, and the latter having higher scores in recall metrics. These process discovery algorithms will be applied to the data within the ProM framework, after which it will have discovered two petri nets. Split-miner is omitted, as it has not been implemented in the ProM framework.

Conformance-checking techniques will be applied with these petri nets. Rozinat et al. have implemented a *Conformance Checker* in the ProM framework, which allows for evaluating if an instance of a process is considered non-conformant [15].

Lastly, with the framework de Leoni et al. introduced into ProM, a decision tree will be created that classifies if a good does or does not have a defect[9].

### 3.5 Evaluation

The evaluation is conducted in three parts, corresponding to the methods described in section 3.4. To evaluate the models the Heuristics and ILP-miner created, the *Multi-Process View Explorer* will be used. This ProM-plugin measures the fitness and precision of petri nets. After this, we will compare the two models based on those metrics.

To determine whether event log traces associated with defects in manufactured goods are non-conformant, logs for both traces linked to goods with defects and goods without defects will be replayed on the petri nets discovered by the Heuristics and ILP-miner, after which they can be classified as conformant or non-conformant, where we will measure those classifications in the metrics precision and recall.

To evaluate the relationship between process features and the quality of manufactured goods, distinguishing between those with and without defects. We will assess the previously modeled decision tree based on its classifications by measuring its precision and recall metrics.

### 3.6 Deployment and Dissemination

This research stops at dissemination by sharing this paper with the company and presenting the results to them. The deployment is left to the company.

## 4 DATASET

This section aims to describe the structure of the datasets and the selection of attributes considered relevant to the research. Two datasets are analyzed. The first one is a synthetic dataset, which is described in section 4.1. The second dataset described in section 4.2 is for a case study of Lienesch B.V. This dataset is data directly pulled from their ERP system.

| | Precision | Fitness |
|---|---|---|
| HeuristicsMiner | 0.560 | 0.307 |
| ILP-miner | 0.085 | 0.809 |

Table 1. Precision and fitness scores for the models discovered by the HeuristicsMiner and ILP-miner on the synthetic dataset, calculated by the Multi-perspective Process Explorer plugin in ProM.

### 4.1 Synthetic Dataset

A synthetic dataset [12] is used to study the methods in a setting that contains less noise first. ERP systems are known for having difficult data to extract, as many features within an ERP system's database exist. The dataset consists of 4544 event logs, which contain activities related to manufacturing processes within a facility. Most importantly, this synthetic dataset contains *rework* values, which will emulate our customer complaints. As the quality is considered not good enough if a manufactured good is rejected, it is possible to test our hypotheses on this dataset first.

### 4.2 ERP-system Dataset

The local organization uses the ERP-system *Exact Globe* [2], of which the database structure is not public. However, the database contains a table that saves event logs; from these event logs, a subset has been exported into a CSV format [17]. With this extraction, the event logs contain all the data starting at the point where the semi-manufactured products come in at the logistics center, go through the production area, followed by a quality check, and finally are stored in the warehouse as a finished product, waiting to be ordered by a customer.

There are 20 692 cases, where 377 (1.82%) have received complaints. Each case has been labeled to indicate whether it has received any complaints.

In this format, the dataset is directly imported into ProM for analysis.

## 5 SYNTHETIC DATASET ANALYTICS

This section describes the results derived from the methodologies described in section 3 applied to the synthetic dataset [12].

### 5.1 Modeling and Evaluation

*5.1.1 Process Discovery.* Process discovery algorithms do not score well on precision and fitness out of the box on most datasets [10]. The dataset considered reflects a real-life manufacturing process with these problems. Applying the HeuristicsMiner and ILP-miner results in the precision and fitness as detailed in table 1. These metric scores are from the models created by transforming the dataset into an XES format and processing it with the Heuristics and ILP-miner.

The following pre-processing methods have been applied to raise the fitness and precision needed for proper data analysis.

*Artificial Sources and Sinks.* For the HeuristicsMiner, determining the starting and termination activities posed a significant challenge, which is needed for the evaluation plugins. To address this, an artificial *Source* and *Sink* activity has been added to every case. After this,

*Filtering per type of manufactured good.* When inspecting the dataset, it's intuitively clear that the process for the creation of different goods differs significantly. This is why the dataset is filtered per good, after which recall and precision scores improve. In a random selection process, cable heads and spur gears were chosen for a detailed analysis. Table 2 shows the recall and precision scores for these processes when split by type of good, highlighting the manufacturing process of cable heads and spur gears.

*5.1.2 Conformance Checking.* Test logs are analyzed by replaying them on the established process models to determine their conformity. Table 3 shows the confusion matrix for this classification concerning the production of cable heads on the process model discovered by the HeuristicsMiner. A non-conformant trace is considered a defect, and a conformant trace is classified as having no defect. Table 4 presents identical data for the process model derived from the ILP-miner.

*5.1.3 Relation between Features-analysis.* Using the *Perform Prediction of Business Process Features* plugin in ProM, it is possible to find which process features have the highest relation to the attribute that describes whether a mistake has been made in the manufacturing process[9]. In the synthetic dataset, that means that there is an annotation that *rework* has been performed. A selection is made of all process features and rework, after which rework is set as the dependent variable. The plugin then calculates the decision tree with the highest F1-score, of which two are shown in figure 1 for cable heads and figure 2 for spur gears. Both parts were randomly selected as parts produced multiple times and had at least one case where the rework flag was set to true. To analyze these decision trees, confusion matrices have been constructed for the cable heads in Table 5 and for the spur gears in Table 6.

*5.1.4 Model Comparison.* Having these confusion matrices, table 7 is constructed. The table shows the precision and recall for classifying defects for the three models. Check for conformity on the process models created by the heuristic and ILP-miner and the decision tree, with the highest score for both metrics given in bold-faced text.

For the objective of identifying the maximum number of defects accurately, the decision tree model stands out due to its high precision. This means it is most effective at correctly identifying defective goods without falsely labeling nondefective goods as defective.

On the other hand, high recall is essential if the goal is to maximize the detection of defects while minimizing the misclassification of nondefective items as defective. Again, the decision tree model has the highest metric, indicating its effectiveness in correctly identifying defects.

Their low fitness severely impacts the recall of both models created by the process discovery algorithms. That metric directly impacts the number of nondefective manufactured goods classified as defective. The low precision of the model also means that traces not present in the event log the model was trained on have a higher chance of being classified as conformant. They are possibly resulting in a lower precision in the classification.

| | Precision (HM) | Fitness (HM) | Precision (ILP) | Fitness (ILP) |
|---|---|---|---|---|
| Cable head | **0.482** | 0.618 | 0.223 | **0.943** |
| Spur gear | **0.762** | 0.416 | 0.178 | **0.921** |

Table 2. Precision and fitness of process models after splitting the event logs by different types of goods and adding artificial sources and sinks for the HeuristicsMiner (HM) and ILP-miner (ILP).


Fig. 1. This decision tree delineates the process feature classification for cable head production, identifying if rework is required ('True') or not ('False'). Each leaf node indicates the outcome ('True/False') along with the classification accuracy in the format (correctly classified/incorrectly classified instances).

| | | True value | | |
|---|---|---|---|---|
| | | No Defects | Defects | Total |
| Classified as | No Defects | 640 | 147 | 787 |
| | Defects | 397 | 91 | 488 |
| | Total | 1037 | 238 | 1275 |

Table 3. Confusion matrix for the classification of conformant and non-conformant traces on the model of creating cable heads produced by the HeuristicsMiner.

| | | True value | | |
|---|---|---|---|---|
| | | No Defects | Defects | Total |
| Classified as | No Defects | 981 | 214 | 1195 |
| | Defects | 56 | 24 | 80 |
| | Total | 1037 | 238 | 1275 |

Table 4. Confusion matrix for the classification of conformant and non-conformant traces on the model of creating cable heads produced by the ILP-miner.

## 6 REAL LIFE DATA ANALYTICS

This section delves into the context of our chosen case study, highlighting Lienesch B.V., a business-to-business curtain manufacturer. The aim is to dissect their manufacturing process, particularly the production of pleated curtains. The objective is to analyze and predict potential defects in fabric pieces based on process features and conformance, thus enhancing quality control measures within the manufacturing process.

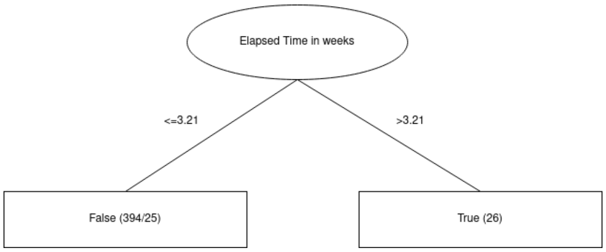


Fig. 2. This decision tree delineates the process feature classification for spur gear production, identifying if rework is required ('True') or not ('False'). Each leaf node indicates the outcome ('True/False') along with the classification accuracy in the format (correctly classified/incorrectly classified instances).

| | | True value | | |
|---|---|---|---|---|
| | | No Defects | Defects | Total |
| Classified as | No Defects | 941 | 20 | 961 |
| | Defects | 96 | 218 | 314 |
| | Total | 1037 | 238 | 1275 |

Table 5. Confusion matrix for cable head defects classification by the decision tree in figure 1.

| | | True value | | |
|---|---|---|---|---|
| | | No Defects | Defects | Total |
| Classified as | No Defects | 368 | 1 | 369 |
| | Defects | 25 | 26 | 51 |
| | Total | 393 | 27 | 420 |

Table 6. Confusion matrix for spur gear defects classification by the decision tree in figure 2.

| | Precision | Recall |
|---|---|---|
| HeuristicsMiner Model | 0.382 | 0.186 |
| ILP-miner Model | 0.101 | 0.300 |
| Decision Tree | **0.916** | **0.694** |

Table 7. Precision and recall metrics for classifying cable heads as defective or non-defective in three models.

## 6.1 Business and Data Understanding

Understanding the business and its data is crucial for our analysis. Figure 3 describes the perceived understanding of the manufacturing process at the local company. This perceived process, or *ground truth* from here on, is created by the understanding of the physical process of making a manufactured good. However, it must be noted that this perceived understanding is not data-driven yet. It is not transparent for all the stakeholders in the organization whose traces in the event logs belong to which part of the process. While the Petri net contains 14 different activity types, the event logs contain 37, to which we conclude that the process is not so trivial as perceived in the ground truth.

| | Precision | Fitness |
|---|---|---|
| HeuristicsMiner | 0.901 | 0.782 |
| ILP-miner | 0.811 | 0.836 |

Table 8. Precision and fitness scores for the models discovered by the HeuristicsMiner and ILP-miner on case study dataset, calculated by the Multi-perspective Process Explorer plugin in ProM.

| | | True value | | |
|---|---|---|---|---|
| | | No Defects | Defects | Total |
| Classified as | No Defects | 15 886 | 126 | 16012 |
| | Defects | 4 429 | 251 | 4680 |
| | Total | 20 315 | 377 | 20 692 |

Table 9. Confusion matrix for classifying conformant and non-conformant traces on the model of creating pleated fabrics produced by the HeuristicsMiner. Non-conformant traces are classified as fabric having defects, and conformant traces are classified as goods with no defects.

## 6.2 Data Preparation

Having understood the business context, we now focus on preparing the data for analysis. In section 5.1, it was noted that process discovery algorithms do not score well out of the box, especially on datasets retrieved from ERP systems. The noise, usually present within an ERP system, has been eliminated by only selecting the relevant activities through the initial SQL query that created our dataset [17]. Unlike the synthetic dataset, the process in the case study started and ended with the same activity for every case, resulting in not having to add artificial sources and sinks.

## 6.3 Modeling and Evaluation

This section's core is applying process mining techniques to the case study data, following the methodology outlined in Section 3.

*6.3.1 Process Discovery.* In our process discovery exploration, we applied the Heuristics and ILP-miners to our dataset. As shown in Table 8, the resulting models reveal that the HeuristicsMiner scores higher in precision on this dataset. In contrast, the ILP-miner scores higher in recall. While an ideal scenario would involve higher scores for both models, time constraints in our research necessitated proceeding with the current results.

*6.3.2 Conformance checking.* Next, we assess the conformance of the manufacturing process through an analysis of the event logs. By replaying traces of event logs of produced items where customers have and have not complained, we classify those goods based on whether a trace was conformant or non-conformant to the process model. If a trace is non-conformant, it is categorized as having a defect, while conformant traces are classified as having no defects. The results are shown in table 9 and 10 for the Heuristics and ILP-miner, respectively.

*6.3.3 Relation between Features-analysis.* The correlation between various process features is our final point of interest for creating a model in our case study. Again, we use the *Perform Prediction of Business Process Features* plugin in ProM, as employed in section

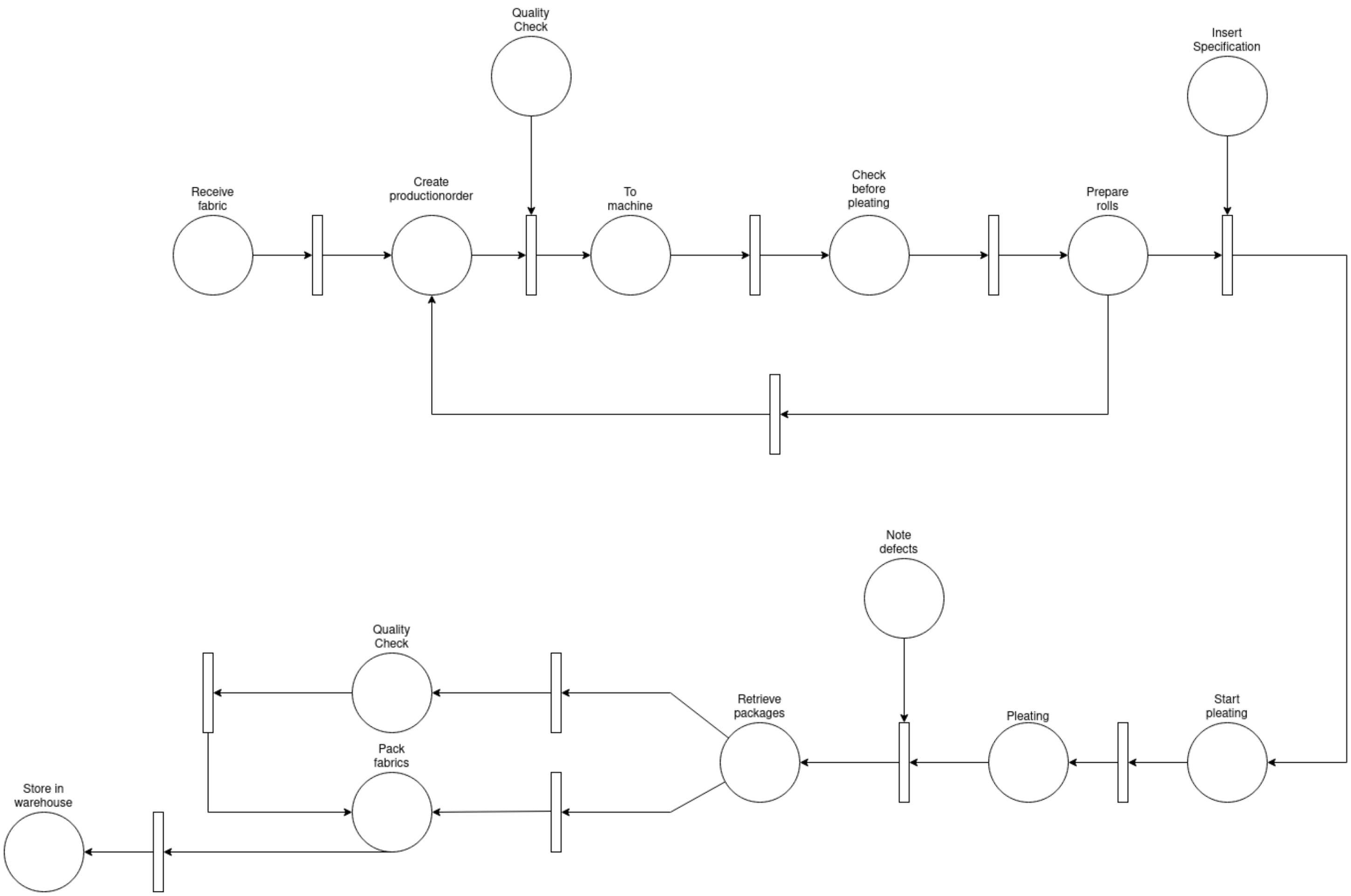


Fig. 3. Perceived understanding of the production process of pleated fabrics at Lienesch B.V. visualized as a Petri net.

| | | True value | | |
|---|---|---|---|---|
| | | No Defects | Defects | Total |
| Classified as | No Defects | 16979 | 190 | 17169 |
| | Defects | 3336 | 187 | 3523 |
| | Total | 20 315 | 377 | 20 692 |

Table 10. Confusion matrix for classifying conformant and non-conformant traces on the model of creating pleated fabrics produced by the ILP-miner. Non-conformant traces are classified as fabric having defects, and conformant traces are classified as goods with no defects.

| | | True value | | |
|---|---|---|---|---|
| | | No Defects | Defects | Total |
| Classified as | No Defects | 20 118 | 156 | 20 274 |
| | Defects | 197 | 221 | 418 |
| | Total | 20 315 | 377 | 20 692 |

Table 11. Confusion matrix for defects in pleated fabrics classification by decision tree in figure 4.

5.1.3. This plugin allows us to relate process features to data points. The decision tree, as shown in figure 4, is the result of the decision tree with the highest F1-score achieved by classifying the goods as having or not having defects. The confusion matrix resulting from said decision tree is portrayed in table 11.

*6.3.4 Model Comparison.* We now analyze which model best suits our needs with the confusion matrices, classifying the event logs based on the model produced by the heuristic miner, ILP-miner, and the decision tree. Table 12 compares the performance of heuristic miner, ILP-miner, and decision tree models using precision and recall metrics. The heuristic miner model excels in precision, while the decision tree leads in recall. For Lienesch B.V., balancing accurately identifying defective goods against the false classification of non-defective ones is crucial so as not to waste resources. Despite the heuristic miner's higher precision, the significant recall advantage of the decision tree model makes it a more suitable choice in this context.

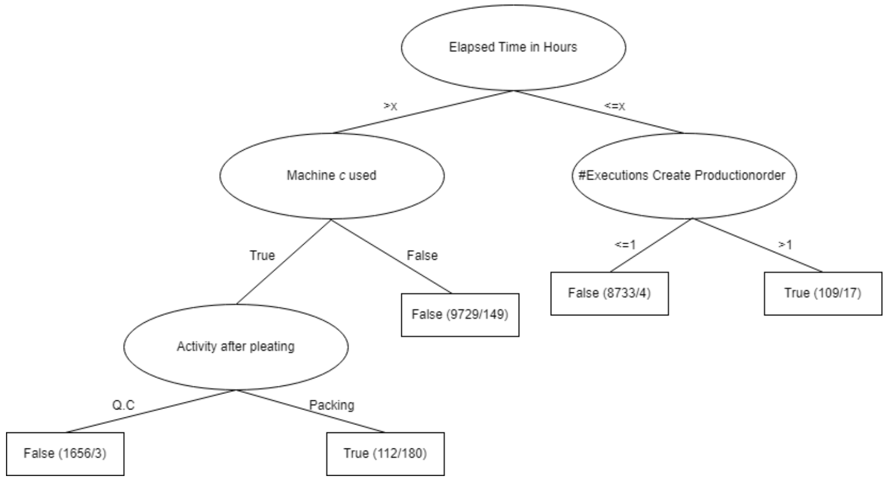


Fig. 4. This decision tree delineates the process feature classification for pleated fabrics production, identifying if the fabric has received complaints ('True') or not ('False'). Each leaf node indicates the outcome ('True/False') along with the classification accuracy in the format (correctly classified/incorrectly classified instances) **N.B.** Some values have been omitted, as the information is not intended to be public.

| | Precision | Recall |
|---|---|---|
| Heuristic Miner Model | **0.666** | 0.054 |
| ILP-miner Model | 0.496 | 0.053 |
| Decision Tree | 0.586 | **0.529** |

Table 12. Precision and recall metrics for classifying pleated fabrics as defective or non-defective in three models.

## 7 CONCLUSION

In this research, we aimed to address the central question posed in the introduction: "How can we develop a predictive model to anticipate complaints about manufactured goods?"

We compared the HeuristicsMiner to the ILP-miner on a synthetic and real-life dataset in section 5 and 6, respectively. This showed that the HeuristicsMiner achieved higher precision process models while the ILP-miner achieved higher fitness process models. Our findings revealed that manufacturing data often lacks the structure for high-precision and high-fitness process models. As a result, many conformant processes were wrongly labeled non-conformant. This is detailed in Table 3 and 4 for a synthetic dataset, and in 9 and 10 for the real-life dataset. Consequently, identifying non-conformant processes is less useful for organizations. This is because non-conformant processes are rare, and verifying cases where processes were incorrectly classified as non-conformant wastes resources. We emphasized identifying critical process features and establishing their relationship with the target data point to overcome this challenge. The result is a decision tree that effectively classifies manufacturing defects, as shown in figure 1 and 4. This approach minimizes the misclassification of defect-free products, ensuring efficient resource utilization.

In summary, our research provides practical insights into creating predictive models for defect anticipation in manufacturing processes:

- For structured processes, conformance checking of traces for defect anticipation is not working well enough to be used in a manufacturing process. Because too many goods without defects are classified as having defects. However, the results are promising enough to re-evaluate when the fitness of models is increased.
- For both structured and less-structured processes, we advise the development of decision tree models that select relevant process features guided by a business understanding. Based on these identified features, this model can effectively classify products as defective or defect-free.

## 8 DISCUSSION

This section addresses the limitations of our research, which serve as the basis of our future work suggestions.

In our study, we confined ourselves to using ProM [18], which limited our process discovery algorithms to be used to the ones implemented in the framework. Consequently, other well-performing algorithms, like Split-Miner [6], were not used. Additionally, an algorithm has been developed to improve the fitness of models generated by the HeuristicsMiner [16]. Implementing this algorithm and reassessing its impact on our research could be beneficial, as it might render the HeuristicsMiner models more competitive compared to the decision tree models, particularly if there is a large enhancement in fitness.

Time limitations constrained our project, particularly in the data pre-processing phase. Pre-processing is crucial for cleaning data to make it more suitable for process mining techniques [13], but this aspect was not extensively addressed in our study.

For future research, it is advisable to construct the *ground truth representation* in a BPMN-model [3] instead of a Petri net [1]. Although the Petri net depicted in figure 3 is believed to accurately represent the process as described by employees of Lienesch B.V., BPMN models are typically more intuitive and require less specialized knowledge for interpretation [14].

Another avenue for future research could involve training models exclusively on datasets comprising goods with defects. If a process aligns with a model based solely on defective goods, it might indicate a higher likelihood of defects in the produced goods.

## ACKNOWLEDGMENTS

Many thanks to my supervisor dr. F.A. Bukhsh for her supervision and the weekly sparring sessions which certainly helped this research. I also want to express my heartfelt gratitude to Lienesch B.V. and its employees for their help and the freedom to perform the research I wanted on their dataset. A big thanks go out to Bob, Liesbeth, Martijn en Ruud in particular.

## REFERENCES

[1] 2022. *Process Mining Handbook*. Springer International Publishing. https://doi.org/10.1007/978-3-031-08848-3
[2] 2024. Exact Globe. https://www.exact.com/nl/software/exact-globe.
[3] Gustav Aagesen and John Krogstie. 2015. BPMN 2.0 for Modeling Business Processes. *Handbook on Business Process Management 1: Introduction, Methods, and Information Systems* (04 2015), 219–250. https://doi.org/10.1007/978-3-642-45100-3_10
[4] Wil Aalst and A. Weijters. 2004. Process mining: A research agenda. *Computers in Industry* 53 (06 2004), 231–244. https://doi.org/10.1016/j.compind.2003.10.001
[5] R. Addo-Tenkorang and Helo P. 2011. Enterprise Resource Planning (ERP): A Review Literature Report. *Proceedings of the World Congress on Engineering and Computer Science 2011 Vol II* (10 2011).
[6] Adriano Augusto, Raffaele Conforti, Marlon Dumas, Marcello La Rosa, and Artem Polyvyanyy. 2018. Split miner: automated discovery of accurate and simple business process models from event logs. *Knowledge and Information Systems* 59, 2 (May 2018), 251–284. https://doi.org/10.1007/s10115-018-1214-x
[7] Martin Bauer, Han van der Aa, and Matthias Weidlich. 2019. Estimating Process Conformance by Trace Sampling and Result Approximation. In *Business Process Management*, Thomas Hildebrandt, Boudewijn F. van Dongen, Maximilian Röglinger, and Jan Mendling (Eds.). Springer International Publishing, Cham, 179–197.
[8] R. P. Jagadeesh Chandra Bose and Wil M.P. van der Aalst. 2009. Context Aware Trace Clustering: Towards Improving Process Mining Results. In *Proceedings of the 2009 SIAM International Conference on Data Mining*. Society for Industrial and Applied Mathematics. https://doi.org/10.1137/1.9781611972795.35
[9] Massimiliano de Leoni, Wil M.P. van der Aalst, and Marcus Dees. 2016. A general process mining framework for correlating, predicting and clustering dynamic behavior based on event logs. *Information Systems* 56 (2016), 235–257. https://doi.org/10.1016/j.is.2015.07.003
[10] Jochen De Weerdt, Manu De Backer, Jan Vanthienen, and Bart Baesens. 2012. A multi-dimensional quality assessment of state-of-the-art process discovery algorithms using real-life event logs. *Information Systems* 37, 7 (2012), 654–676. https://doi.org/10.1016/j.is.2012.02.004
[11] Sebastian Dunzer, Matthias Stierle, Martin Matzner, and Stephan Baier. 2019. Conformance Checking: A State-of-the-Art Literature Review. In *Proceedings of the 11th International Conference on Subject-Oriented Business Process Management* (Seville, Spain) *(S-BPM ONE '19)*. Association for Computing Machinery, New York, NY, USA, Article 4, 10 pages. https://doi.org/10.1145/3329007.3329014
[12] Dafna Levy. 2014. Production Analysis with Process Mining Technology. https://doi.org/10.4121/UUID:68726926-5AC5-4FAB-B873-EE76EA412399
[13] Heidy M. Marin-Castro and Edgar Tello-Leal. 2021. Event Log Preprocessing for Process Mining: A Review. *Applied Sciences* 11, 22 (Nov. 2021), 10556. https://doi.org/10.3390/app112210556
[14] Ivo Raedts, Marija Petkovic, Yaroslav Usenko, Jan Martijn Van der Werf, Jan Groote, and Lou Somers. 2007. Transformation of BPMN Models for Behaviour Analysis. 126–137.
[15] A. Rozinat and W.M.P. van der Aalst. 2008. Conformance checking of processes based on monitoring real behavior. *Information Systems* 33, 1 (2008), 64–95. https://doi.org/10.1016/j.is.2007.07.001
[16] Yaguang Sun and Bernhard Bauer. 2016. A Novel Heuristic Method for Improving the Fitness of Mined Business Process Models, Vol. 9936. 537–546. https://doi.org/10.1007/978-3-319-46295-0_33
[17] Darrell Tufto. 2023. Production Event Logs Lienesch B.V. Data extracted from Lienesch B.V. ERP system.
[18] Eric Verbeek. 2024. ProM framework. https://github.com/promworkbench.
[19] R. Wirth and Jochen Hipp. 2000. CRISP-DM: Towards a standard process model for data mining. *Proceedings of the 4th International Conference on the Practical Applications of Knowledge Discovery and Data Mining* (01 2000).
[20] Jing Yang, Shaobo Li, Zheng Wang, Hao Dong, Jun Wang, and Shihao Tang. 2020. Using Deep Learning to Detect Defects in Manufacturing: A Comprehensive Survey and Current Challenges. *Materials* 13, 24 (2020). https://doi.org/10.3390/ma13245755