Machine Learning-Supported Planning of Lead Times in Job Shop Manufacturing

Kathrin Julia Kramer, Carsten Wagner, Matthias Schmidt

Machine Learning-Supported Planning of Lead Times in Job Shop Manufacturing

Kathrin Julia Kramer[1], Carsten Wagner[2] and Matthias Schmidt[1]

[1] PPI, Leuphana University Lueneburg, 21335 Lueneburg, Germany
[2] HAWK University of Applied Sciences and Art Hildesheim/ Holzminden/ Goettingen, Buesgenweg 1a, 37077 Goettingen, Germany
kathrin.kramer@leuphana.de, carsten.wagner@hawk.de, matthias.schmidt@leuphana.de

Abstract. In order to ensure adherence to schedules, knowledge of planned lead times (LT) is crucial for success. In practice, however, rigid planning methods are often used which cannot adequately reflect constantly changing environmental influences (e.g. fluctuations in the daily workload). Particularly in job shop production, precise planning of LT is difficult to implement. This paper therefore examines whether existing machine learning (ML) approaches, in particular supervised learning methods, in production planning can support LT scheduling in job shop production to generate added value. The paper enhances existing research by comparing deep artificial neural networks with ensemble methods (e.g. random forest, boosting decision trees). The applied approach bases on the Cross Industry Standard Process for Data Mining (CRISP-DM), which was created by a consortium of companies. Finally, the evaluation through an exemplary job shop production shows that the present work contributes to mastering the planned LT. In particular, the ML model, boosting decision trees and deep artificial neural networks show significant improvements in planning quality. This practical reference has not yet been addressed comprehensively in the literature.

Keywords: Machine Learning, Production Planning & Control Approaches, Job Shop Production, Lead Times.

1 Introduction

In production companies, the lead time (LT), inventory, capacity utilization and adherence to delivery dates are the production logistical objectives. However, it is not possible to improve all objectives at the same time, as some of them are contradictory. For example, high capacity utilization can be achieved through high stocks, which in turn results in long LT. Long LT usually fluctuate strongly and thus have a negative influence on adherence to delivery dates [1, 2, 3]. The mutual influence underlines the importance of precise planning of the respective logistical objectives.

In times of increasing customer requirements and thus a high level of adherence to delivery dates, realistic planned LT are necessary in order to be able to confirm a feasible date to the customer. The correct determination of planned LT is also important for the determination of capacities within the production system and for procurement [4]. For instance, as Shaw and Whinston [5] outline, more precise LT have a positive effect on the accuracy of scheduling decisions and vice versa. Incorrect LT planning can lead to low adherence to delivery dates and high inventories. In practice, however, rigid methods are often used that cannot adequately reflect constantly changing environmental influences (e.g. short-term sick leave of employees or fluctuations in daily workload). Classically used methods often base their prediction on historical data or estimations. Generally, none or limited data is being used in relation to these methods and the actual status of the production system is not reflected. For instance, the average order LT, the sum of the average LT of operations or the sum of the average execution times of the relevant operations of the orders plus a lump sum of the transition time of these operations is used for LT prediction of a new order [1, 2]. In addition, precise planning of LT is particularly difficult in job shop production, because the length of orders, the use and sequence of workstations per order varies: In a job shop a job follows an individual and predefined sequence or route for its processing on one or more existing machines [1, 3]. Machine learning (ML) methods could potentially improve prediction quality, as research has outlined its potential benefit for production planning and control tasks [6]. Therefore, this paper examines whether existing ML methods can provide benefit to the planning of LT in job shop manufacturing.

The structure of the paper foresees four further sections: Section 2 provides an overview about ML in general and the current state of research concerning the prediction of LT with ML methods in a complex production system. Section 3 presents the methodology for applying ML to LT planning and section 4 deploys the process to real data of a job shop manufacturing company. The last section outlines the key findings and a possible future research agenda.

2 Current State of Research

2.1 Definition of ML and its Types

Statements by Arthur Samuel [7] define ML as the field of study in which humans do not dictate every step to the computer and instead facilitate computers to learn independently. One possible way of clustering ML methods is according to their learning method. Following Russel et al. [8], commonly known learning methods are supervised learning (SL), unsupervised learning and reinforcement learning. This paper focuses on regression methods as part of SL methods, thus a function should be identified that predicts LT of orders based on a labelled data set. The model learns to recognize and generalize the relationship between given input and output data and uses it to predict outputs for unknown examples.

2.2 LT Prediction Supported by ML Methods

The use of regression methods for the planning of LT has been carried out using case studies of different production environments. For complex production systems like job shop manufacturing, researchers have explored the use of different input factors, ML methods and data sources. So far, a generalized model selection does not exist and depends on the production environment and the taken model parameters [9]. Hence, it is not suitable to focus on just one model. Instead, a comparison of different models should take place. The current research has used linear regression (LR), decision tree (DT), random forest (RF), k-nearest neighbor and support vector regression (SVR). Further lasso regression, ridge regression, artificial neural networks (NN), multivariate adaptive regression, deep neural networks (DNN), bagging decision trees and boosting decision trees (BDT) were used for determining planned lead times of orders. In regards, to a suitable choice of a regression model, the case studies show that more complex regression models (e.g. RF or BDT) often outperform simple regression models (e.g. LR or DT) in regards to prediction accuracy. However, within the field of more complex regression models, the current literature does not compare DNN with ensemble methods (e.g. RF, BDT).

Gyulai et al. [9] compare the prediction accuracy of LR, DT, RF and SVR using the normalized root mean square error (NRMSE). The modelling is based on data of a real flow-shop environment in the optic industry, where complexity of the LT prediction exists due to a large number of process parameters and uncertainties within foreseeing the client order stream. Input factors are e.g. time of order's arrival in the production system, product type or material. The RF model shows the best performance with a NRMSE that is in average 10.1 percentage points lower than the other models [9]. Lingitz et al. [10] explore the prediction accuracy of LR, lasso regression, ridge regression, NN and multivariate adaptive regression. Further, they examine SVR, k-nearest neighbor, DT, RF, bagging decision trees and BDT. The models use real data (e.g. work in progress of process step or weekday of order's arrival time) of a job shop manufacturing from the semiconductor industry. The RF model and the BDT show the best performance using evaluation metrics like root mean square error (RMSE) or NRMSE [10]. Wang et al. [11] examine the prediction of completion times and not LT by using NN and DNN. The models use real data like the work in progress of each process step within a complex job shop from an equipment manufacturing enterprise. The applied evaluation metrics like RMSE or mean value of the relative errors show that DNN outperforms NN [11].

In contrast to the presented research, this paper aims at covering all aspects of a data mining process to predict LT in a comprehensible way as Cadavid et al. [12] underpins this as essential. A comparison of the state-of-the-art in regards to model selection takes place. This is applied to a real job shop, since simulation data is not suitable for making general assumptions for a real production environment [13].

3 Research Methodology

The research methodology follows the Cross Industry Standard Process for Data Mining (CRISP-DM), a widespread, industry-independent life cycle process [14]. The process contains six phases: "business understanding", "data understanding", "data preparation", "modelling", "evaluation" and "deployment" [14]. This chapter explains each step of CRISP-DM in the context of predicting LT in a job shop environment.

The business-understanding phase covers the set-up of the project by determining business and data mining goals as well as an assessment of the current situation in order to define the scope of the project [14]. As described in the previous chapters, the goal is to examine if a regression model could improve the prediction accuracy of planned LT in a job shop-manufacturing environment. In addition, the complexity of the models should also be taken into consideration. The data-understanding phase foresees to collect initial data and then to describe, examine and to check the data quality [14]. According to Kuepper [15] data can be collected through an analysis of documents or the organisation. Further, stated by Loos [16], data types to be captured are master data or variable data, which can be analyzed e.g. by visualization or statistics [17]. The data preparation phase chooses the relevant data for the modelling phase and prepares the data by cleaning, adding further data or integrating different data sources as well as adjusting the data format [14]. Feature selection or the creation of more suitable features can occur e.g. with regression methods or expert insights [12]. Ludwig & Nyhuis [3] state that the consideration of the order-specific situation by characterizing the orders, recording the current work in progress and capacity situation of the production system and the processing sequence is relevant for the LT prediction. In general, LT consists of the components transition time and execution time. Transition time corresponds to the transportation time, post- and pre-processing waiting time. Execution time is equivalent to the set-up and processing time [1, 2, 18]. Typically, the transition time, especially the waiting time, reflects up to 90% of the overall LT in a job shop [1]. Therefore, especially the present state of influencing factors from the waiting time should contribute to the accuracy of the prediction model. According to Nyhuis [2, 19] the current work in progress at each workstation has a major impact on transition time.

During the modelling phase, chosen regression methods are modelled and a generated test design evaluates the models [14]. Based on the current research LR, DT, RF, BDT and DNN are chosen. DNN is based on a multi-layer feed-forward artificial neural network that is using back-propagation [20]. On the one hand, RF, BDT and DNN show in different studies the best performance, despite a direct comparison between RF/BDT and DNN. On the other hand, LR and DT are less complex, hence easier to understand and implement [8]. The final tradeoff between model complexity and prediction accuracy depends on the business and data mining goals. The test design consists of 40% from the overall data set and measures with RMSE, a commonly used evaluation metric [21]. The tool RapidMiner [20] is applied due to its straightforwardness application of ML models [12].

The evaluation phase evaluates the results, conducts a review process and determines the next steps in regards to a potential deployment [14]. Hence, this phase compares the results of the modelling phase with the originally set goals of the business-understanding phase. The evaluation of the results can either lead to adjustments of the previous phase and therefore to a repetition of the modelling or to an approval of an appropriate model. In the case of an approval, the deployment phase begins by planning the deployment and runtime management as well as by preparing a final project report and conducting a project assessment [14]. For deployment, e.g. considerations of database requirements or the implementation in existing processes need to be conducted. Generally, the characteristics of the manufacturing system or process can change over time

[22]. Therefore, a ML model needs to be updated regularly, as the relationships between the features can change unforeseeable [23]. Hereafter, CRISP-DM corresponds to a continuous improvement process [14].

4 Application

The research methodology is applied to manufacturing enterprise system data from a German job shop manufacturer, who is responsible for the maintenance of complex investment goods. The job shop consists of 14 shop sections, each with three to six workstations of the same kind and has nine main production routes with different work contents and recursion loops. Typically, the order characteristics vary for each maintenance order because the complex capital goods are exposed to different external factors such as environmental conditions or degree of capacity utilization, resulting in different maintenance efforts. Hence, maintenance orders usually differ in their duration and the workstations they pass through, which is consistent with the definition of a job shop. Currently, the planned LT equals to the operation-specific execution time plus a generalized transition time. In an investigation period of about one year, the current method deviates on average by -12.17 days with a standard deviation of 12.95 days from the actual LT. The typical variance of the characteristics from the maintenance orders in a job shop and the prediction deviations of the current method underpin the need for an alternative approach to determine the planned LT. For the modelling phase, thus, the initial data set with the maximum available 66 features describing these orders is examined. The provided data set of approx. 7,900 orders from the ERP system is of sufficient quality and consists of features, which characterize the orders (e.g. business unit, client type, part type, order quantity, and planned working hours per workstation), the order's process (e.g. actual start of processing) or the job shop (e.g. work in progress per workstation). The feedback was accurate to the day, which is sufficient for predicting LT of orders. Statistical and visual analysis do not show strong patterns for deviating LT or any outliers. However, the data set excludes 13 attributes, because they are not available at the time of forecasting a new order's LT (e.g. actual start of processing) or they reflect the planned working hours of a workstation, which is not used in the majority of orders and thus has insufficient quality of information for the prediction. Further feature selection is conducted through an algorithm-based approach, whereas the choice also depends on the set parameters of the models. Referring to table 1, the models all have the originally planned LT as one of their top features, which ranges from 1.6 to 233.9 days with an average of 27.6 days and a standard deviation of 10.1 days. In four of the five models, the planned working hours of the workstation 131 is a major feature. Each order uses this workstation with varying planned working hours of 0.10 to 78.00 hours with an average of 2.89 hours and a standard deviation of 4.11 hours. The workstations 81, 91, 21 and the overall planned working hours are only in some of the chosen models a major factor. Not amongst the top three features, nevertheless the work in progress of some specific workstations has also an influencing effect and underpins Nyhuis's hypothesis [2, 19].

As stated in table 1, a comparison of the RMSE values from the different approaches shows that all chosen models outperform the original procedure and DNN has overall the best accuracy. However, the method is only slightly better than BDT. The RMSE of LR, DT and RF is roughly one day higher than from BDT and DNN. Overall, the RMSE of the DNN model, which is not as traceable as simpler models, is currently 9.6 days lower than the original procedure and 1.1 days lower than the LR model. In addition, DNN deviates in average by 0.43 days with a standard deviation of 7.98 days from the actual LT.

Table 1. Overview of results from LT planning.

Model	Top 3 features	RMSE [days]
LR	Originally planned LT, planned working hours of workstation 131 and 81	5.9
DT	Originally planned LT, overall planned working hours, planned working hours of workstation 131	6.0
RF	Originally planned LT, planned working hours of workstation 131 and 91	6.1
BDT	Originally planned LT, overall planned working hours, planned working hours of workstation 91	5.0
DNN	Originally planned LT, planned working hours of workstation 131 and 21	4.8
Original	Operation-specific execution time, generalized transition time	14.4

Moreover, figure 1 visualizes the model accuracy by showing the distribution of deviations between actual and planned LT from the test data set of LR, DNN and the original procedure.

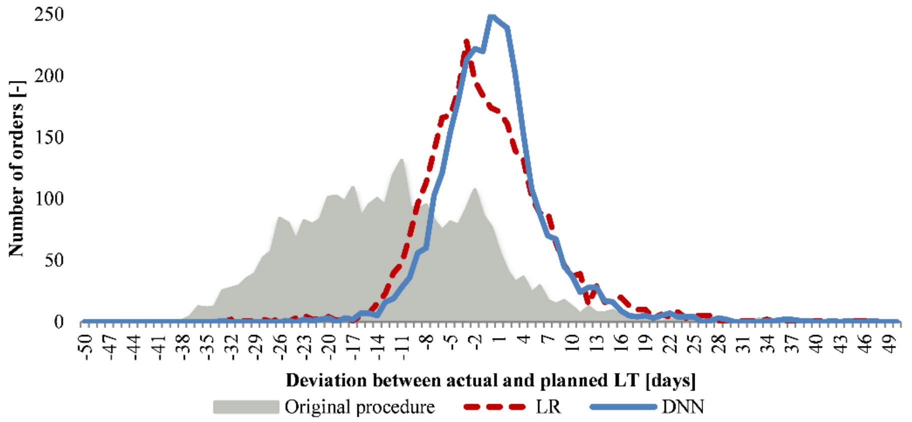

Fig. 1. Deviation between actual and planned LT of different prediction models

The average deviation between actual and planned LT of the test data set has decreased by 11.83 days for LR and 12.61 days for DNN in comparison to the original procedure. In order to decrease the overall range of deviations, a further balancing of the data set could be carried out, as currently the majority of orders consist of short LT. In addition, as Ludwig & Nyhuis [3] state, information about the processing sequence of operations could further add value to the accuracy of the models.

Within the scope of a feasibility study of the company, the applicability was examined and evaluated based on actual data. Due to the high programming effort in the ERP system used, an approach has not yet been integrated into the ongoing operation.

5 Conclusion and Future Research Agenda

This paper has shown a comprehensive approach to predict LT in a job shop environment. Overall, more precise LT planning reduces the effort of downstream process steps (e.g. capacity coordination) in job shop production and enables a more reliable promise of customer appointments. The paper contributes to the current research by highlighting the different state of the art of regression model selection and the associated input characteristics. The application to a real job shop production shows that DNN has the best prediction accuracy, closely followed by BDT. RF's accuracy is slightly worse than the simpler models, LR and DT. However, in current literature RF has often shown the best results, which cannot be proven for this case.

Further investigations should be carried out in an additional research project to identify generally valid statements along the modelling process for LT prediction. The aim is to study the different influence factors in LT planning and to identify the key influence variables. This would also facilitate the introduction of SL approaches, as correspondingly targeted priorities could be set within the data-mining project. The research project could also be further specified by additional analyses of other regression models or modifications with regard to parameter setting. Besides the research about ML-supported LT planning, there is also a need for further research on the general handling of ML in the production domain. For example, practically relevant research questions concern ethics, data protection, and the traceability of the solution or the integration of ML models into existing organizational structures. In a further step, these findings can be applied to the determination of planned LT.

Acknowledgements. Founded by the Lower Saxony Ministry of Science and Culture under grant number ZN3489 within the Lower Saxony "Vorab" of the Volkswagen Foundation and supported by the Center for Digital Innovations (ZDIN).

References

1. Wiendahl, H.-P.: Betriebsorganisation für Ingenieure. 8th edn. Carl Hanser, Muenchen (2014).
2. Nyhuis, P., Wiendahl, H.-P.: Logistic Production Operating Curves – Basic Model of the Theory of Logistic Operating Curves. CIRP Annals 55(1), pp. 441-444 (2006).
3. Ludwig, E., Nyhuis, P.: Verbesserung der Termineinhaltung in komplexen Fertigungsbereichen durch einen neuen Ansatz zur Plandurchlaufzeitermittlung. In: Görke, W., Rininsland,

H., Syrbe, M. (eds.), Information als Produktionsfaktor, pp. 473-475. Springer, Heidelberg (1992).

4. Lödding, H.: Handbook of Manufacturing Control – Fundamentals, Description, Configuration. Springer, Berlin Heidelberg (2016).
5. Shaw, M., Whinston, A. B.: An Artificial Intelligence Approach to the Scheduling of Flexible Manufacturing Systems. IIE Transactions, 21 (2), pp. 170-183 (1989).
6. Bullers, W. I., Nof, S. Y., Whinston, A.B.,: Artificial Intelligence in Manufacturing Planning and Control. AIIE Transactions, 12 (4), pp. 351-363 (1980).
7. Samuel, A. L.: Some Studies in Machine Learning Using the Game of Checkers. IBM Journal of research and development 3(3), pp. 210-229 (1959).
8. Russel, S., Norvig, P.: Artificial Intelligence: A modern approach. 3rd edn. Pearson Education, Upper Saddle River (2010).
9. Gyulai, D., Pfeiffer, A., Nick, G., Gallina, V., Sihn, W., Monostori, L.: Lead time prediction in a flow-shop environment with analytical and machine learning approaches. IFAC-PapersOnLine 51(11), pp. 1029-1034 (2018).
10. Lingitz, L., Gallina, V., Ansari, F., Gyulai, D., Pfeiffer, A., Sihn, W., Monostori, L.: Lead time prediction using machine learning algorithms: A case study by a semiconductor manufacturer. 51st CIRP Conference on Manufacturing Systems 72, pp. 1051-1056 (2018).
11. Wang, C., Jiang, P.: Deep neural networks based order completion time prediction by using real-time job shop RFID data. Journal of Intell Manuf 30(3), pp. 1303-1318 (2019).
12. Cadavid, J. P. U., 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. Journal of Intell Manuf (2020).
13. Mourtzis, D., Doukas M., Fragou, K., Efthymiou, K., Matzorou V.: Knowledge-based estimation of manufacturing lead time for complex engineered-to-order products. 47th CIRP Conference on Manufacturing Systems 17, pp. 499-504 (2014).
14. Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., Wirth, R.: CRISP-DM 1.0 Step-by-step data mining guide. The CRISP-DM consortium (2000).
15. Kuepper, H.-U.: Controlling. Konzeption, Aufgaben und Instrumente. 4th edn. Schaeffer-Poeschel, Stuttgart (2005).
16. Loos, P.: Grunddatenverwaltung und Betriebsdatenerfassung als Basis der Produktionsplanung und -steuerung. In: Corsten, H., Friedl, B. (eds.), Einführung in das Produktionscontrolling, pp- 227-252. Vahlen, München (1999).
17. Witten, I., Frank, E., Hall, M.: Data Mining – Practical Machine Learning Tools and Techniques. 3rd edn. Elsevier, Burlington (2011).
18. Bechte, W.: Methoden und Hilfsmittel der Durchlaufzeit- und Bestandsanalyse in Klein- und Mittelbetrieben. Beuth, Berlin (1979).
19. Nyhuis, P.: Practical Applications of Logistic Operating Curves. CIRP Annals 56(1), pp. 483-486 (2007).
20. Mierswa, I., Klinkenberg, R.: RapidMiner Studio. 9.3. RapidMiner Inc. (2019).
21. Kuhn, M., Johnson, K.: Applied Predictive Modeling. 1st edn. Springer, New York (2013).
22. Deep, K., Singh, P. K.: Design of robust cellular manufacturing system for dynamicpart population considering multiple processing routes using geneticalgorithm. Journal of Manufacturing Systems 35, pp. 155-163 (2015).
23. Hammami, Z., Mouelhi, W., Said, L. B.: On-line self-adaptive framework for tailoring a neural-agent learning model addressing dynamic real-time scheduling problems. Journal of Manufacturing Systems 45, pp. 97-108 (2017).