Potentials of Explainable Predictions of Order Picking Times in Industrial Production

Kaja Balzereit[a], Nehal Soni[b] and Andreas Bunte[c]
Fraunhofer IOSB, Industrial Automation branch, Fraunhofer Center for Machine Learning, Lemgo, Germany

Keywords: Explainability, Prediction, Regression Analysis, Industrial Application of AI.

Abstract: The order picking process in a manufacturing supermarket is central in many industrial productions as it ensures that the items required for production are provided at the right time. However, the order picking process itself often is a black box, i.e., the time it takes to pick an order and the dependencies in the process that influence the time usually are not exactly known. In this work, we highlight the potentials of creating explainable predictions of order picking times using Artificial Intelligence methods. The prediction is based on the analysis of a historic database and on a linear regression analysis that learns the dependencies in the data. From this prediction, (1) the potential of identifying features having a high and a low influence on the order picking time, (2) the potential of optimizing the order picking process itself, and (3) the potential of optimizing depending processes are identified. For prediction, we utilize the regression methods LASSO and Decision Tree. These methods are compared with regard to their interpretability and usability in industrial manufacturing.

1 INTRODUCTION

Artificial Intelligence (AI) gets more and more important in many areas of life (European Factories of the Future Research Association, 2019). Due to the ability of AI of handling and analyzing large data with a high accuracy, its potential is huge (Burkart and Huber, 2021). However, for AI to be applicable to sensitive domains, for example when it comes to processing data about human behavior, it needs to be *explainable*, i.e., the reasoning steps of the AI need to be comprehensible for a human. But methods from Machine Learning (ML) often comprise many mathematical transformations and aggregations that prevent an intuitive understanding. Explainable AI (XAI) allows for deeper insights into AI models, and thus, allows for drawing conclusions about the underlying, unknown analyses (Balzereit et al., 2022).

Whereas explainability is crucial for AI in domains such as medicine (Holzinger et al., 2019), it is also required for the domain of manufacturing. AI harbors a great potential throughout the whole industrial production (Lu, 2019; World Intellectual Property Organization, 2019). And modern computing infrastructures enable the collection of data about every step of production. Nowadays, manufacturing is a human-centered, cyber-physical process (Monostori, 2014). Hence, the data that is stored about manufacturing processes undoubtedly involves data about human operators. So for an AI system to be applicable to manufacturing processes, its use and analysis of data shall be understandable. Furthermore, many manufacturing processes are black-box processes, i.e., the dependencies between different influencing factors usually are not known to its fullest extent (Balzereit et al., 2019). XAI allows for drawing conclusions about these dependencies, and thus, gives deeper insights into unknown processes.

In this article, we are concerned with the application of XAI to the order picking process in a manufacturing supermarket. A manufacturing supermarket is a decentralized warehouse storage that allows for quick and easy access to items required for production. The principle of a manufacturing supermarket goes back to the Toyota Production System: a picker collects those items required for production of a specified product from easily accessible shelves (Towill, 2010). The shelves are restocked as soon as required with items from a less accessible storage, e.g., a high-rack storage (Yang et al., 2015). The process of col-

[a] https://orcid.org/0000-0001-9203-5902
[b] https://orcid.org/0000-0002-3149-2971
[c] https://orcid.org/0000-0001-6878-0419

Balzereit, K., Soni, N. and Bunte, A.
Potentials of Explainable Predictions of Order Picking Times in Industrial Production.
DOI: 10.5220/0011677000003393
In *Proceedings of the 15th International Conference on Agents and Artificial Intelligence (ICAART 2023) - Volume 3*, pages 405-412
ISBN: 978-989-758-623-1; ISSN: 2184-433X

lecting items required is called order picking process.

The order picking process itself usually is a black-box, i.e., the time required for picking one order is not exactly known. Hence, in production planning, which needs to incorporate many depending processes, order picking time often is not considered. As nowadays in production lot sizes tend to shrink and delivery times shall be minimized, it can be error prone to not consider order picking times (European Factories of the Future Research Association, 2019).

For optimizing the order picking process and connected processes, the order picking time needs to be estimated from available data. But dependencies between KPIs, such as picking times and input features, are not known (Balzereit et al., 2019; Balzereit et al., 2022). In this article, we apply XAI to an industrial order picking process. Our contribution is as follows:

(1) We will show an intuitive approach to predict order picking times of new orders, based on historic data and expert knowledge. For this purpose, features ensuring interpretability are generated from expert knowledge. Interpretable regression methods ensure a directly interpretable model.

(2) We compare LASSO regression and decision tree regression towards their usability and interpretability in predicting order picking times in manufacturing supermarkets.

(3) We outline the potential of predicting order picking times for some AI applications in production planning: (a) The used models allow for extracting information about the influence of features. Thus, information for adapting the process to reduce order picking times can be drawn. (b) A prediction of order picking times for single orders allows for optimizing the sequence of orders. Thus, load peaks can be reduced and a homogeneous load factor of the pickers can be reached. (c) Processes subsequent to order picking can be optimized by integrating an estimation of the order picking time.

This article is structured as follows: first, the related work is described (Section 2). Then, in Section 3, the order picking process is described. Our approach on predicting the order picking time is presented in Section 4. In Section 5, the results are presented. After that, the potentials of the prediction are highlighted (Section 6). Finally, a conclusion is given in Section 7.

2 RELATED WORK

Dovsilovic et al. (Došilović et al., 2018) understand XAI as a pedagogical system that enables humans to understand the reasoning of complex algorithms. This process is crucial for the wide application of AI. Furthermore, as AI need to be continuously improved, either by updating the training data or by adjusting the hyperparameters of the model, a deep understanding of the model and how it was created is essential (Ahmed et al., 2022).

Nor et al. (Nor et al., 2021) emphasized the potential of XAI in industrial applications such as prognostics, diagnostics, and anomaly detection. It states that the research interest in XAI in industrial applications rises continuously. Especially interpretable models, rule- and knowledge-based approaches, and attention mechanisms, which enable understanding of image recognition, face a steep rise of interest.

Burkart and Huber (Burkart and Huber, 2021) published an exhaustive article about recent advantages in XAI. They classified approaches in *interpretable by nature*, *interpretable by design*, and *black-box models*. Approaches that are interpretable by nature are not optimized in a special way to achieve interpretability but interpretability is intrinsic in the approach. For example, linear regression, least absolute shrinkage and selection operator (LASSO), or the CART algorithm that is used to train decision trees are interpretable by nature. Interpretability by design refers to models for which the interpretability can be controlled. Examples are Deep Neural Decision Trees, that combine decision trees with neural networks (Yang et al., 2015), and Ordered Rules for Classification, an approach that creates a sequence of decision rules that are identified using Mixed Integer Optimization. Black-box approaches, in general, comprise all methods that do not come with an explanation of their decision. For example, deep neural networks are black-box approaches as their decision is based on the concatenation of mathematical transformations on the input data. An explanation of why the specific prediction is given cannot be extracted easily. To achieve explainability for black-box models, post-hoc methods creating an explanation after a black-box model has been trained, are needed.

Related work from the area of order picking is concerned the use of convolutional neural networks for classification (Grzeszick et al., 2017) and on the use of different optimization techniques to increase the efficiency of the process (Ene and Öztürk, 2012; Moeller, 2011). Hence, the potential in the analysis of data about the order picking process is huge.

Our approach uses AI methods that are interpretable by nature. We target an intuitive understanding and a gain of insights into the industrial manufacturing process. Thus, hidden dependencies in the data are identified.

3 THE ORDER PICKING PROCESS

Order picking is an essential and cost demanding process in the supply chain in industrial production. Given a list of items required for manufacturing a product, a picker walks through a manufacturing supermarket to collect these items. The layout of the supermarket, the picking system, and the storage and routing strategies are important aspects that affect the efficiency of the order picking process (Burinskiene, 2010).

The order picking data is collected as the picker navigates through the warehouse to collect the items. Every time a picker starts a new order, finishes an order, or picks an item, a scanner documents the particular process step. Each of these operations is assigned with a timestamp. When picking an item, along with the timestamp, the weight of the item and the location of the item in the supermarket is stored. In addition, for each picking order the type of trucks required for transportation of the items and the lot size (amount of order) is recorded. This data constitutes the raw data of the order picking process.

For example, manufacturing an engine requires an enclosure, a coil and a rotor among other items. While a single order requires only one quantity of each item, an order of four engines requires four quantities of each item which can be collected in a batch. The shelf trucks are used for smaller items, while pallet trucks are used for larger items. For heavier items, the picking time can be increased as a picker needs to lift it and the truck gets heavier.

To understand the different aspects affecting the order picking time, we use an ML technique to process raw data for analyzing and predicting picking time for future orders.

4 EXPLICIT PREDICTION OF ORDER PICKING TIME

In this work, we use supervised, interpretable-by-nature AI methods. Supervised algorithms use features (input) and target (output) data to model the process. An ML model is a linear or non-linear mathematical representation of the process dynamics. Features are information contributing to the outcome that we want to predict, i.e., target. The target variable extracted from the historic data is referred to as *ground truth*. The goal of supervised ML modeling is to reliably predict the target variable when new unseen data is given to the model.

Fig. 1 represents our approach in general. Our goal is to predict the time required to pick an order. For this purpose, first, we extract the relevant information from the historic raw data of this process. Factors such as the layout of the manufacturing supermarket, weights of products, the means to carry them around, and others contribute to the time it takes to pick an order.

Our approach comprises two steps: first, features are extracted from the raw data (Section 4.1). Then, a regression is fitted, learning the dependency between the features and the order picking time (Section 4.2).

4.1 Feature Engineering

Identifying useful features from the raw data is a crucial part of ML modeling. An example of the features calculated in our use case is given in the *Feature Engineering* block of Fig. 1.

The target *order picking time* is calculated by taking a difference of timestamps (end time and start time) of the whole process of picking one order.

To predict this target, various features, which supposedly influence the order picking time, are created. The *number of items* is the number of different items required for production of a product. The *weight of items* is the sum of individual items' weight. *Types of truck* refers to the vehicle being used to carry the product depending on where it is located on a shelf. The *storage types* is added as a feature to understand its impact on time. In many manufacturing supermarkets, storage spaces are separated into fixed and flexible spaces, i.e., a fixed space stores always the same item whereas the items stored in a flexible space may vary with different days. The *lot size* refers to a number of the same items to be picked.

We have also incorporated expert knowledge in calculating certain features. For an instance, the raw data contains information about the specific location where an item is placed in the supermarket. Combined with an encoded layout of the supermarket, this data is used to calculate an *approximated distance* a picker has to cover to collect all the items. This distance calculation is carried out such that the shortest distance will be considered — to resemble human behavior. Further expert knowledge such as the impact of different types of truck usage on picking time enabled us to utilize them as features. This expert knowledge enables (i) to include all features which presumably have an impact on the order picking time and (ii) to calculate features which are directly interpretable. Automatically generated features, in contrast, require no expert knowledge for the generation but also comprise features with no or a hard practical

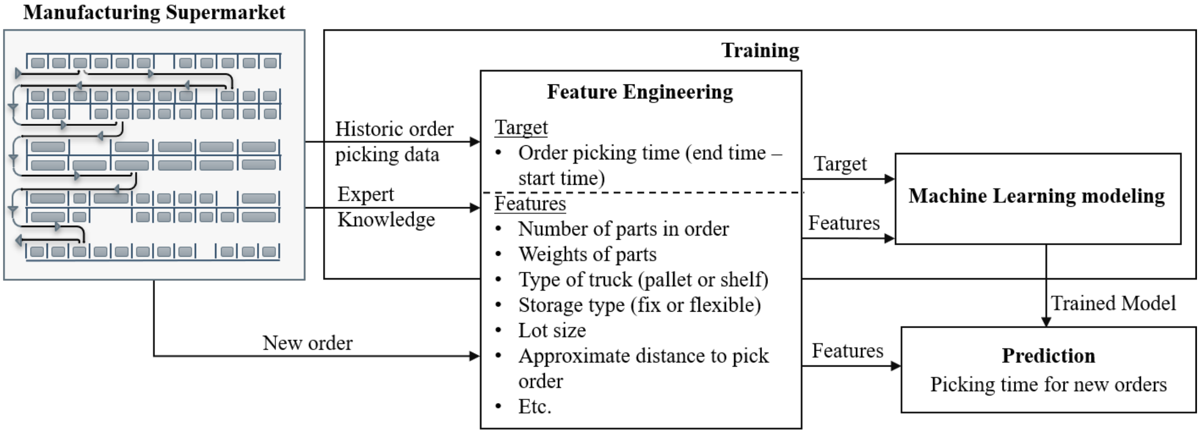

Figure 1: Prediction process using feature engineering and machine learning modeling.

interpretation.

4.2 Training of Regression Model

Our target variable, order picking time, is a continuous value. Hence, a prediction can be achieved by regression.

Here, we utilize LASSO Regression (Section 4.2.1) and Decision Tree Regression (Section 4.2.2). Both algorithms, in a first step, given a historic database, learn a dependency between the input features and the target time. This step is also called training phase. In a second step, also called operational phase, the learned model is used to calculate a prediction for a new order, which is not in the historic database.

The model performance is evaluated using various metrics such as Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), R^2 score. Especially the RMSE and the MAE are directly interpretable, as their unit is the same as the target variable. R^2 (R-squared), a commonly used metric for linear regression, represents how much of the variation in the target variable can be explained by taking features into the account.

4.2.1 Linear Regression

LR generates the best fitting line which calculates the output variable by summing up the weighted features. The feature weight, also known as *coefficient*, is parameterized by $w_i \in \mathbb{R}$ for feature $x_i \in \mathbb{R}, i \in \{1,2,...n\}$. $n \in \mathbb{N}$ represents the number of features and w_0 is the intercept of the fitting line. So, a prediction of target variable y can be calculated as

$$y = \sum_{i=1}^{n} w_i \cdot x_i + w_0. \tag{1}$$

LR models a dependency between independent and dependent variables by minimizing the residual sum of squares, so

$$\min \frac{1}{N} \sum_{j=1}^{N} (y_j - \hat{y}_j)^2 \tag{2}$$

where $N \in \mathbb{N}$ is the number of observations, y is the prediction of the linear regression and $\hat{y}$ is the ground truth value. Hence, it is also called *least squares Linear Regression*. However, when the number of features is high, LR tends to overfit, i.e., the coefficients values are fitted too strong to the training data and the transferability to the test data is reduced. Hence, we add a least absolute shrinkage and selection operator (LASSO). LASSO linear regression adds the squared sum of coefficients to the minimization function, i.e.,

$$\min \frac{1}{N} \sum_{j=1}^{N} (y_j - \hat{y}_j)^2 + \alpha \sum_{i=0}^{n} w_i^2, \tag{3}$$

where $\alpha \in \mathbb{R}_{\geq 0}$ is a regularization factor. Thus, coefficients intentionally are held as small as possible. Coefficients of features having a negligible impact on the target are set to zero.

4.2.2 Decision Tree Regression

A Decision Tree partitions the feature space using bound constraints, e.g., $x_{i_1} \leq t_{i_1}$ with $t_{i_1} \in \mathbb{R}$ (Hastie et al., 2009) (see Figure 2 for an example). Starting at a root node, constraints on the feature space create two new branches - one if the constraint is satisfied and one if not. Thus, a tree of decisions is created. At the end of each decision path, a leaf node assigns the examined order a prediction value.

There are various algorithms for learning a tree. The most popular one is presumably CART (Breiman et al., 2017) which aims at maximizing the information gain. However, many efficient algorithms, e.g.,

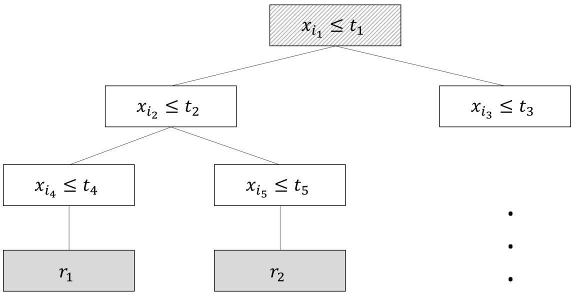

Figure 2: Example of a Decision Tree. The root node is patterned, leaf nodes are filled gray.

ID4.5 (Utgoff, 1989), have been published in the last years.

4.3 Operational Phase

In the operational phase, the learned model is used to predict the order picking time of new orders. For this purpose, for the new order, also the features are extracted using the same feature generation procedure used for training. These features then are given as an input to the learned model.

5 RESULTS

5.1 Application Case

The examined supermarket contains around 300 places for different parts, from which 200 are fixed places and 100 are flexible places. A typical order comprises between ten and 20 different parts, typical lot sizes are between 8 and 30.

The typical time for one process is between four and twelve minutes.

5.2 Performance

We performed two different types of regression: Lasso regression and Decision Tree Regression. The results of the linear regression are illustrated in Figure 3; the results of the decision tree regression are illustrated in Figure 4. Black points represent the ground truth, i.e., the actually measured time for commissioning stored in the test data. Blue points represent the prediction of the corresponding regression method.

The mean absolute error for the linear regression is approximately 2.4 minutes; the mean absolute error for the decision tree regression is approximately 2.5 minutes.

From the regression, it can be drawn that the features *lot size* and *number of different products in an order* are very relevant for the prediction of the time required for commissioning.

5.3 Comparison of Decision Tree Regression and LASSO Regression for Explainability

Table 1 discusses LASSO and Decision Tree Regression in terms of interpretability and applicability. In general, both approaches are interpretable in training and prediction. LASSO regression minimizes squared residuals, whereas DTR creates a tree consisting of constraints on features.

DTR allows for the direct extraction of decision rules, which, according to our estimation, are more easily to understand for operators who have few prior knowledge about AI. LASSO regression creates an equation allowing for estimating the quantitative influence of features on the target variable. This equation, however, requires some knowledge about mathematics or AI, respectively, to allow for an appropriate interpretation. Furthermore, as regression requires a normalization of features a-priori, further transformations are required to allow for an interpretation of coefficients in accordance with the un-normalized order of magnitude of the feature.

However, LASSO creates a continuously differentiable equation representing the dependency between the features and the target variable. This equation can directly be used in a subsequent optimization step. As modern optimization algorithms mainly operate on gradient descent methods, the property of continuous differentiability enables the application of many optimization algorithms. DTR, on the other hand, creates a sequence of decision rules, that, however, is not continuously differentiable. Hence, further manual efforts are required to integrate such a prediction into an optimization method.

The number of features may also pose a challenge for AI algorithms. As LASSO contains a term that reduces the magnitude of coefficients, also many features can be handled well. Features which a small importance actively are assigned a zero coefficient. Decision Trees, however, are well known to tend to overfitting. Here, a careful parametrization of the learning algorithm is necessary to enable a proper handling in case of many feature available.

6 DISCUSSION

Predicting order picking times harbors diverse potentials for further AI applications. In this section, var-

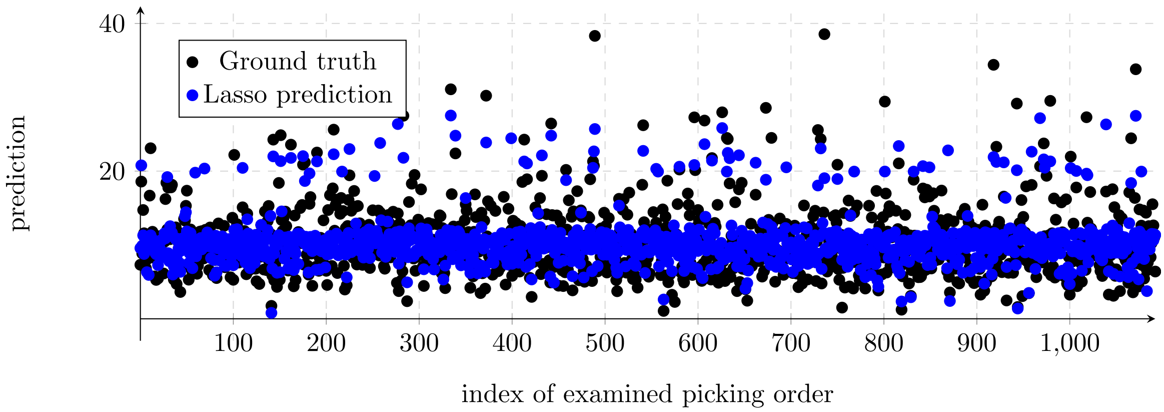

Figure 3: Prediction results for LASSO.

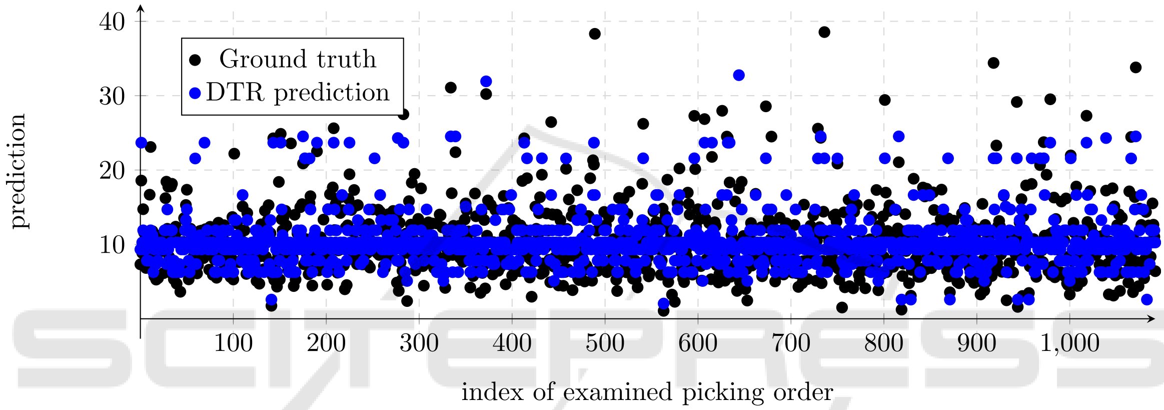

Figure 4: Prediction results for decision tree regression.

ious potentials are discussed. For reducing the order picking times, knowledge about features having a large influence on the order picking time is crucial. In Section 6.1, how to extract these features is presented. The sequence of picking orders has potential for optimization, as it may ensures deadlines to be met and buffer times to be reduced. A discussion of this AI application is given in Section 6.2. The optimization of other processes which depend on the order picking process is presented in Section 6.3.

6.1 Identification of Features with Low and High Influence

From LASSO Linear Regression, the influence of each feature can be directly drawn from the learned model (3). So, previously unknown dependencies, e.g. that the total weight of a picking order has a severe impact on the picking time, can be extracted. This information about the influence of different features on the picking time allows for identifying adaptations to the picking process that reduce the picking time. For example, if the total weight of a picking order has a sever impact on the picking time, the picking orders can be adapted such that the total weight is reduced, e.g. by splitting large orders into multiple single orders. In addition, features that have a very small or negligible impact can be identified as their coefficients are either close to zero or zero.

From the decision tree, decision rules directly can be extracted: each path from the root node to a leaf node creates a set of intuitively interpretable decision rules. These rules consist of constraints on the feature. After application of all decision rules extracted from a path, a decision tree assigns a value to the examined picking order. Furthermore, decision trees allow for drawing conclusion about the importance of individual features. Features in nodes close to the root cause usually have a higher influence on the decision than features closer to leave nodes. However, the absolute frequency of a feature also gives information about the importance of a feature (Tierney et al., 2022).

Thus, practitioners are enabled to evaluate different adaptations that aim at reducing the time of the picking process. Using the knowledge about the im-

Table 1: Comparison of LASSO Regression and Decision Tree Regression in terms of explainability and applicability for industrial manufacturing.

Criterion	LASSO	DTR
General interpretability of learned model	yes	yes
Extraction of decision rules	no	yes
Direct integration in optimization algorithm possible	yes	no
Appropriate handling of many features	yes	no

portance of the features, new rules aiming at reducing the specific features can be created. For example, in case the weight has a crucial impact on the picking time, reducing the total weight of orders could lead to a reduction of time of the picking process.

6.2 Optimization of Order Picking Process

The main goal of modeling the picking process is its optimization, by using this data-driven approach. Typically, companies have a stock of orders that have to be processed, until a certain point in time. Therefore, pickers get the next order on the list and start the picking process. Since we can predict the time of picking orders, we can use the model to optimize that. Sequence planning is well known in production (SAP,) and can now be transferred to the order picking process. The optimization criteria can be manifold, e.g., increasing the efficiency, shifting workloads from night-shift to day-shift, increasing the buffer for the production, or avoiding workload picks. In the following, we will describe two criteria in more detail, shifting workloads and increasing efficiency.

During the order picking process, the highest priority is to provide always enough material for the production. To optimize the shifts, the predicted number of workers in every shift should be close to a whole number, to avoid underemployed workers. Furthermore, orders with a high degree of picking effort and low production effort can be planned during the day shift, whereas orders with low picking effort and high production effort can be completed during the night shift. This enables to reduce the number of required pickers at night.

To increase the efficiency of the picking process, the sequence of orders can be adjusted or orders can be combined. Changing the sequence is relatively easy to implement, but depending on the lot sizes the effect might be small. Combining orders with small lot sizes has more potential, since picking larger lot sizes is generally more efficient. However, many constraints have to be considered, e.g. the capacity of the trucks, the delivery time for each product, and the load factor of each line, if multiple production lines are provided. Anyway, if all constraints are known, they can be considered by planning. We suppose that we can achieve good results with heuristics, which can process large order sets in short times.

6.3 Integrating the Order Picking Process into Production

As described above, it is possible to optimize the order picking process by using the model presented in this work. However, the information about how long a picking process takes can be used for optimization of the overall production planning. So, the picking orders can be triggered by the production, which will lead to an optimal production or the production is triggered by available orders, which leads to an optimal picking process. Both will lead to a local optimum of one process. To reach a global optimum, both processes have to be combined in one plan. Besides increased efficiency, it also enables a more customized or reduced buffer, since the processes are more connected and better synchronized. However, the combined optimization increases the solution space, which makes it more challenging to find an optimal solution. This makes it hard or impossible to solve these issues with classical approaches. But since we have a model, new approaches such as reinforcement learning can be used and help to tackle this challenge (Panzer et al., 2021). However, only such a holistic view can enable a more efficient process. From a practical point of view, the creation of an appropriate problem model and the implementation of a suitable solution algorithm are the key challenges here.

7 CONCLUSION

The order picking process in a manufacturing supermarket is the process of a picker collecting items required for production of a specific product. Nowadays, this process often is a black box, so the time needed for collecting all items for a specific order is unknown in advance. Also, the dependencies in the process, e.g., how the number of items that need to be collected affects the order picking time, usually are not known. However, much data about the order picking process is collected, allowing for thorough analyses. In this work, an approach for analyzing this data to generate a prediction of order picking times is presented. The approach is based on explainable AI methods that learn the dependencies between the features of historic picking orders and the time required. A comparison of LASSO and Decision Tree Regression discusses the benefits and drawbacks of the individual approaches. We highlight the potentials this prediction harbors for industrial production: it enables the identification of features affecting the order picking time significantly or just marginally, it allows for optimization of the order picking process itself, and it enables the optimization of overall production as the prediction of order picking time can be integrated into global optimization approaches.

ACKNOWLEDGEMENTS

This work was funded by research grant “005-2001-0031” of the MWIDE NRW in form of the Spitzencluster it’s OWL intitiative in the research project IMAGINE. We thank the Wilo SE for providng the application case.

REFERENCES

Ahmed, I., Jeon, G., and Piccialli, F. (2022). From artificial intelligence to explainable artificial intelligence in industry 4.0: a survey on what, how, and where. *IEEE Transactions on Industrial Informatics*, 18(8):5031–5042.

Balzereit, K., Diedrich, A., Kubus, D., Ginster, J., and Bunte, A. (2022). Generating causal hypotheses for explaining black-box industrial processes. *5th IEEE International Conference on Industrial Cyber-Physical Systems (ICPS)*.

Balzereit, K., Maier, A., Barig, B., Hutschenreuther, T., and Niggemann, O. (2019). Data-driven identification of causal dependencies in cyber-physical production systems. In *11th International Conference on Agents and Artificial Intelligence*.

Breiman, L., Friedman, J. H., Olshen, R. A., and Stone, C. J. (2017). *Classification and regression trees*. Routledge.

Burinskiene, A. (2010). Order picking process at warehouses. *International Journal of Logistics Systems and Management - Int J Logist Syst Manag*, 6.

Burkart, N. and Huber, M. F. (2021). A survey on the explainability of supervised machine learning. *Journal of Artificial Intelligence Research*, 70:245–317.

Došilović, F. K., Brčić, M., and Hlupić, N. (2018). Explainable artificial intelligence: A survey. In *2018 41st International convention on information and communication technology, electronics and microelectronics (MIPRO)*, pages 0210–0215. IEEE.

Ene, S. and Öztürk, N. (2012). Storage location assignment and order picking optimization in the automotive industry. *The international journal of advanced manufacturing technology*, 60(5):787–797.

European Factories of the Future Research Association (2019). EFFRA Vision for a manufacturing partnership in Horizon Europe. 2021-2027.

Grzeszick, R., Lenk, J. M., Rueda, F. M., Fink, G. A., Feldhorst, S., and Ten Hompel, M. (2017). Deep neural network based human activity recognition for the order picking process. In *Proceedings of the 4th international Workshop on Sensor-based Activity Recognition and Interaction*, pages 1–6.

Hastie, T., Tibshirani, R., Friedman, J. H., and Friedman, J. H. (2009). *The elements of statistical learning: data mining, inference, and prediction*, volume 2. Springer.

Holzinger, A., Langs, G., Denk, H., Zatloukal, K., and Müller, H. (2019). Causability and explainability of artificial intelligence in medicine. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 9(4):e1312.

Lu, Y. (2019). Artificial intelligence: a survey on evolution, models, applications and future trends. *Journal of Management Analytics*, 6(1):1–29.

Moeller, K. (2011). Increasing warehouse order picking performance by sequence optimization. *Procedia-Social and Behavioral Sciences*, 20:177–185.

Monostori, L. (2014). Cyber-physical production systems: Roots, expectations and r&d challenges. *Procedia Cirp*, 17:9–13.

Nor, A. K. M., Pedapati, S. R., Muhammad, M., and Leiva, V. (2021). Overview of explainable artificial intelligence for prognostic and health management of industrial assets based on preferred reporting items for systematic reviews and meta-analyses. *Sensors*, 21(23):8020.

Panzer, M., Bender, B., and Gronau, N. (2021). Deep reinforcement learning in production planning and control: A systematic literature review. In *Conference on Production Systems and Logistics (CPSL)*.

SAP. Sequence planning - documentation of sap manufacturing exection.

Tierney, K., Balzereit, K., Bunte, A., and Niehörster, O. (2022). Explaining solutions to multi-stage stochastic optimization problems to decision makers. In *2022 IEEE 27th International Conference on Emerging Technologies and Factory Automation (ETFA)*, pages 1–4.

Towill, D. R. (2010). Industrial engineering the toyota production system. *Journal of Management History*.

Utgoff, P. E. (1989). Incremental induction of decision trees. *Machine learning*, 4(2):161–186.

World Intellectual Property Organization (2019). WIPO technology trends. Artificial Intelligence.

Yang, T., Kuo, Y., Su, C.-T., and Hou, C.-L. (2015). Lean production system design for fishing net manufacturing using lean principles and simulation optimization. *Journal of Manufacturing Systems*, 34:66–73.