

Article

Assembly Line Overall Equipment Effectiveness (OEE) Prediction from Human Estimation to Supervised Machine Learning

Péter Dobra ^{1,*} and János Jósmai ²¹ Doctoral School of Multidisciplinary Engineering Sciences, Széchenyi István University, 9026 Győr, Hungary² Department of Vehicle Manufacturing, Széchenyi István University, 9026 Győr, Hungary; josvai@ga.sze.hu

* Correspondence: dobra.peter@sze.hu

Abstract: Nowadays, in the domain of production logistics, one of the most complex planning processes is the accurate forecasting of production and assembly efficiency. In industrial companies, Overall Equipment Effectiveness (OEE) is one of the most common used efficiency measures at semi-automatic assembly lines. Proper estimation supports the right use of resources and more accurate and cost-effective delivery to the customers. This paper presents the prediction of OEE by comparing human prediction with one of the techniques of supervised machine learning through a real-life example. In addition to descriptive statistics, takt time-based decision trees are applied and the target-oriented OEE prediction model is presented. This concept takes into account recent data and assembly line targets with different weights. Using the model, the value of OEE can be predicted with an accuracy of within 1% on a weekly basis, four weeks in advance.

Keywords: OEE; assembly line; machine learning; prediction; decision tree

**Citation:** Dobra, P.; Jósmai, J.

Assembly Line Overall Equipment Effectiveness (OEE) Prediction from Human Estimation to Supervised Machine Learning. *J. Manuf. Mater. Process.* **2022**, *6*, 59. <https://doi.org/10.3390/jmmp6030059>

Academic Editor: Steven Y. Liang

Received: 19 April 2022

Accepted: 26 May 2022

Published: 27 May 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

For most industrial companies, the prediction of various data is based on computer-aided human decisions and experiences. Production data collection, recording and processing within Industry 4.0 and Smart Manufacturing is today a routine activity [1–6]. Numerous systems, among others, the Manufacturing Execution System (MES), Enterprise Resource Planning (ERP) and Customer Relationship Management (CRM), support data management and forecasting [7–9]. It is essential to be aware of the efficiency of production units in order to properly plan and assemble the sequence of products. It is inevitable to predict the change in efficiency, which can move in either a negative or positive direction in addition to the stagnant situation. Lower efficiency requires more manpower, and more manpower generates higher variable costs. Therefore, it is important to estimate future efficiency values as accurately as possible.

The aim of this paper is to explore, analyze and compare the prediction supported by human and decision trees and reveal a new model for more accurate efficiency prediction.

2. Literature Review

Analyzing historical data is no longer a complex task, although production efficiency is affected by a number of plannable and unplannable factors. The stochastic nature of the human work (e.g., variable assembly time, cycle time, etc.) and the occurrence of the duration of machine faults is particularly difficult in the case of semi-automatic assembly lines [10].

Nowadays, within Artificial Intelligence (AI), the benefits of Data Mining (DM) and Machine Learning (ML) are increasingly being exploited in the field of production [11–15]. Different techniques and algorithms are used in the fault detection, quality improvement, process characterization, predictive maintenance, energy consumption, production planning, cycle time prediction, lead time estimation, and scheduling and decision support system [16–22].

In the domain of machine learning, there are several methods such as: deep learning (e.g., deep belief networks, convolutional neural networks, etc.), ensemble (e.g., random forest, gradient boosting machines, etc.), neural networks (e.g., perceptron, back-propagation, etc.), Bayesian (e.g., Bayesian network, Gaussian naïve Bayes, etc.), regularization (e.g., least absolute shrinkage and selected operator, elastic net, etc.), regression (e.g., linear regression, logistic regression, etc.), dimensionality reduction (multidimensional scaling, linear discriminant analysis, etc.), instant-based (e.g., k-nearest neighbor, learning vector quantization, etc.), clustering (e.g., k-means, hierarchical clustering, etc.), rule systems (e.g., one rule, zero rule, etc.) and decision tree [23–25].

At the domain of the decision tree, there are also several methods, among others, the Classification and Regression Tree (CART), C4.5, C5.0, Chi-squared automatic interaction detection, decision stump, conditional decision tree, M5 and interactive dichotomiser 3 (ID3) [26]. For decision tree models, the main components used are nodes (root node, internal node, leaf nodes), branches (if-then rules), splitting, stopping and pruning.

All machine learning methods can be categorized into the following three groups:

- Supervised Learning (SL);
- Unsupervised Learning (UL);
- Reinforcement Learning (RL) [27,28].

Based on Choudhary et al., the prediction goal is focusing on predicting the behavior of a model and determining the future values based on existing information [29].

OEE as a best practice and standard firstly was introduced within the Total Productive Maintenance (TPM) framework by Nakajima [30]. The original formula for calculation OEE is written as:

$$\text{OEE} = a p q \quad (1)$$

where,

a—availability [%],
p—performance [%],
q—quality [%].

According to Corrales et al., in the scientific literature more than 850 papers were published between 1996 and 2020 about Overall Equipment Effectiveness [31]. Most articles deal with increasing OEE using a variety of methods. This paper seeks to explore how to achieve an accuracy of within 1% for OEE prediction.

3. Materials and Methods

In the field of production logistics, one of the most complex planning processes is the accurate forecasting of production and assembly efficiency. Full customer service is only possible if the forecast quantity can be economically produced and delivered in the specified time. To do this, however, the appropriate resources must be available. In addition, setting manufacturing and assembly targets can only be realistic if it is well defined what the production system will be capable of in terms of output.

According to the original formula, the expected availability is greater than 90%, performance is greater than 95%, and quality is greater than 99% [30]. It follows that in industrial practice, the percentages achieved are the highest in the field of quality, followed by performance, and finally, availability. The higher the target and performance, the fewer deviations, errors and outages are allowed. Considering the factors of availability, performance and quality, the issue of factors influencing OEE can be approached from several perspectives. The value of OEE follows logically from most of the parameters listed, but sometimes the direction of change of each factor requires more careful consideration. In the six big losses approach, reducing the number and duration of the changeover, machine downtime, cycle time at bottleneck station, minor stoppage, rework and scrap rate and startup losses, resulted that availability, performance and quality will increase, thus OEE percentage will be higher [32,33]. Figure 1 shows the approach of six big losses and its effects.

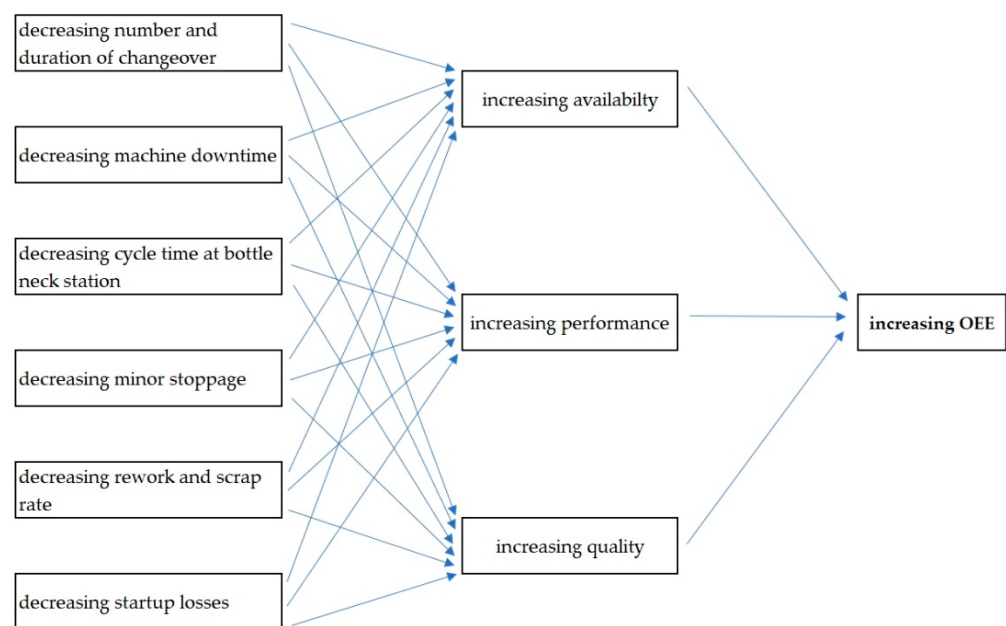


Figure 1. Approach of six big losses and its affects.

In terms of production environment, the value of availability, performance and quality will increase (including OEE) in case of the increasing series size, progress in product life cycle, increasing MTBF, decreasing MTTR, decreasing material shortage and increasing ordered quantity. However, the percentage of OEE increases with higher automation, but in this case, compared to manual assembly, the value of availability decreases. Industry experience has demonstrated that the efficiency of a manual assembly is generally higher than the efficiency of a machine assembly. Productivity, on the other hand, is the opposite. A real industrial example is the assembly of an airbag cable, where the manual assembly efficiency is around 95% and for semi-automatic assembly is around 90%; fully automatic assembly is around 85% for the same product. Some production environment factor and its affects are shown in Table 1.

Table 1. Factors of production environment.

Factor	Availability	Performance	Quality	OEE
Increasing series size	Increasing	Increasing	Increasing	Increasing
Higher automation	Decreasing	Increasing	Increasing	Increasing
Progress in product life cycle	Increasing	Increasing	Increasing	Increasing
Increasing MTBF	Increasing	Steady/increasing	Steady/increasing	Increasing
Decreasing MTTR	Increasing	Steady/increasing	Steady/increasing	Increasing
Decreasing material shortage	Increasing	Increasing	Increasing	Increasing
Increasing ordered quantity	Increasing	Increasing	Increasing	Increasing

Implementing numerous methods, such as different lean techniques, also increases the value of OEE [34–37]. Proper application of engineering methods, e.g., the Failure Method and Effect Analysis (FMEA), will also result in higher OEE [38]. Higher OEE can also be achieved by increasing the number of shifts in production, however, by increasing the length of shifts (from 8 h to 12 h), the higher OEE percentage will not be realized (Table 2).

Heutmann and Schmitt used disturbance detection and prediction as a part of the control concept for a synchronized individual production control. The consequences of the predicted negative effects have been examined and suggestions made to avoid them [39]. Lingitz et al. presents a machine learning algorithm for lead time prediction at a semiconductor manufacturer and tested several models for higher estimation values [40]. Subramaniyan et al. suggested a data-driven algorithm based on MES to predict throughput

bottlenecks in a production system based on active periods of the machines [41]. Anusha and Umasankar present a concept for major factors affecting OEE and set up a model for OEE prediction. Through R studio, 80% of data are taken for the training data set and 20% of data are taken for the testing data set. It was established that a simple moving average is incapable of recording sudden changes in data, but Holt's double exponential smoothing method can achieve the minimum error [42]. Schuh et al. analyzed deviations between planned and real manufacturing processes and evaluated the forecast reliability of production planning. A method was used to determine four relevant deviations as a workstation of a process, date of a process, duration of a process and sequence of the processed order [43]. According to Brunelli et al., a deep learning-based approach was proposed for predicting the future values of manufacturing performance. The Temporal Convolutional Network (TCN) and Long Short-Term Memory (LSTM) concept were applied for estimation [44]. Other machine learning tools also applied for prediction of OEE, such as the Support Vector Machine (SVM), Random Forest by Hassani et al. [45] and derivative Integral Proportional PID machine learning algorithm with SVM by Acosta et al. [46].

Table 2. Methods and organization.

Factor	Availability	Performance	Quality	OEE
Applying Lean methods	Increasing	Increasing	Increasing	Increasing
Decreasing the value of RPN in PFMEA (Risk Priority Number, Process Failure Mode and Effects Analysis)	Increasing	Increasing	Increasing	Increasing
Increasing operation control, management support	Increasing	Increasing	Increasing	Increasing
Increasing length of shifts 8 vs. 12 h	Increasing	Decreasing	Decreasing	Decreasing
Increasing number of shifts 1 vs. 2 vs. 3 shifts	Increasing	Increasing	Increasing	Increasing

The operation of a production system can be assessed in numerous ways. The most common method in industrial practice is to use Key Performance Indicators (KPIs). Within the KPI system, the monitoring of OEE best reflects the change in the value of efficiency [47]. Beyond monitoring, forecasting is also paramount importance. In addition to mathematical and statistical methods, there are several other concepts for estimating, including the Forecasting Management Competence (FMC) [48]. Lee et al. used big data and cloud-based predictive analytics within a Prognostics and Health Management (PHM) approach [49].

4. Results of Prediction

4.1. OEE Estimation by Human

In addition to the framework of Industry 4.0 and IoT tools, human decisions and experiences still play an important role in practical life. In several cases, production planners question the data and forecasts provided by the machines, because the reliability of the support systems is sometimes not satisfactory enough. A real-life example of this is how customer order data recorded and processed in Enterprise Resource Planning (ERP) can change due to the customer's modified needs. Sales coordinators and production planners know which sales orders can be realistic by the given deadlines and which orders will surely be postponed. Taking these factors into account, operational production management can be more accurate and balanced. With the optimal batch size, production can be more economical and efficient. Human is thus still a key factor in the planning and implementation of assembly efficiency.

Between October 2021 and January 2022 (week 43/2021–week 5/2022), a survey was conducted at a European automotive company to predict OEE by humans. Four machine adjusters working in different shifts with several years of manufacturing experience were asked to predict the OEE value for the next shift. For easier prediction, machine adjusters

did not have to predict the OEE value directly, but their estimated number of units produced. From the number of pieces thus given, the predicted OEE value could be calculated according to the following formula:

$$OEE_p = \frac{P_p}{P_t} \quad (2)$$

where,

OEE_p: predicted OEE (%),

P_p: predicted assembly products (pcs),

P_t: theoretical assembly products (pcs).

In addition to the two shifts' operations, the assembly conditions were characterized by Table 3.

Table 3. Assembly conditions.

Assembly Line	Shifts per Day	Theoretical Output	OEE Target	Number of Workstations	Percentage of Machine Assembly
Line 1—hybrid line	2	850 pcs/shift	90%	18	33.3%
Line 2—hybrid line	2	520 pcs/shift	82%	16	31.2%

Aspects considered in human prediction:

- What type of product will be assembled in a given period (next shift);
- How many type changes are planned;
- Ability and practice of available operators (workers with routine and new operators), especially in critical workplaces and bottleneck station(s);
- Suitable raw material is available for continuous assembly operations;
- Reasons for production downtime in the previous short period (e.g., shift, day), as mainly quality defects and machine failures.

Each machine adjuster estimated the expected yield based on the listed aspects. For assembly line 1 and assembly line 2, the predicted and realized OEE values are shown in Figures 2 and 3 as a function of the cases.

The focus of the survey was primarily on how accurate the prediction was and what factors caused each discrepancy. What is most likely to affect the variation and cannot be accurately predicted? The results are depicted in the descriptive statistics shown in Table 4.

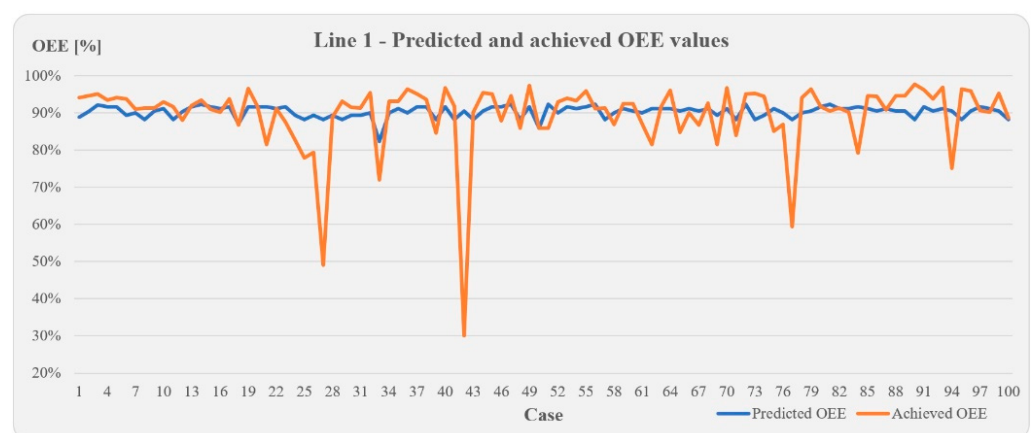


Figure 2. Predicted and achieved OEE values for assembly line 1.

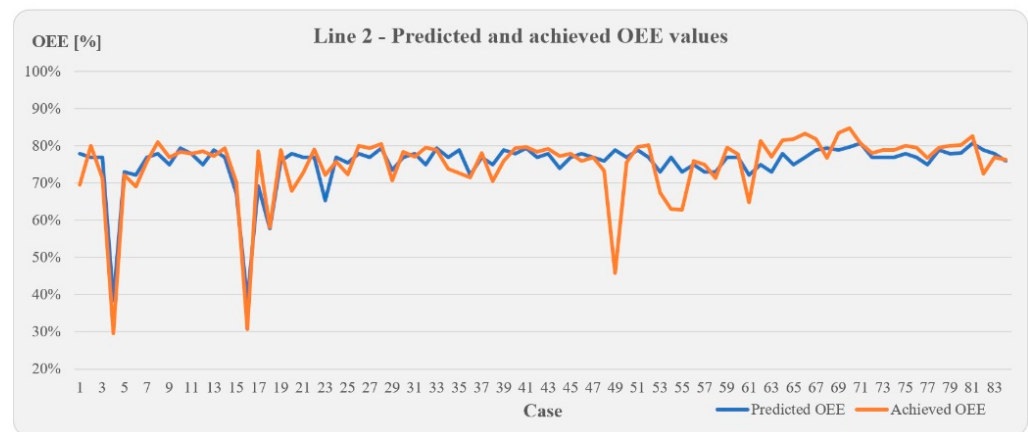


Figure 3. Predicted and achieved OEE values for assembly line 2.

Table 4. Descriptive statistics OEE values for assembly line 1 and assembly line 2.

	Assembly Line 1		Assembly Line 2	
	Predicted	Achieved	Predicted	Achieved
Sample size	100	100	84	84
Minimum	82.4%	30.0%	38.5%	29.6%
Maximum	92.4%	97.8%	80.8%	84.8%
Mean	90.4%	89.6%	75.3%	74.9%
Range	10.0%	67.8%	42.3%	55.2%
Standard error	0.16%	0.94%	0.73%	1.00%
Standard deviation	1.6%	9.4%	6.7%	9.2%

The predicted OEE values move in a relatively narrow range around the target, which means that based on the available data, machine operators do not anticipate a larger change that reduces the OEE value and considers the assembly process to be stable. Nevertheless, there have been several cases where the realized value has differed significantly from the prediction. Most of these discrepancies are negative discrepancies where unexpected and unforeseen events have occurred to varying degrees. The deviation from the estimation is shown in Figures 4 and 5.

The reasons for the TOP 3 negative differences for the two assembly lines are shown in Table 5.

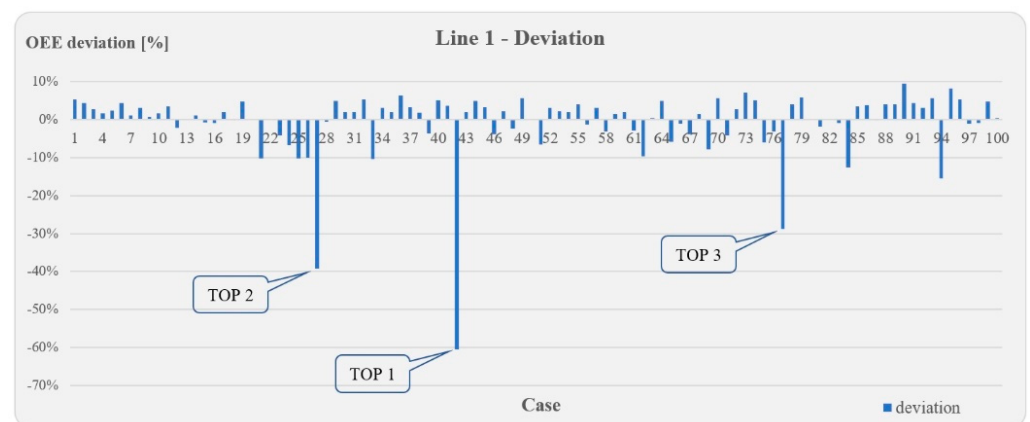


Figure 4. Deviation from the estimation for assembly line 1.

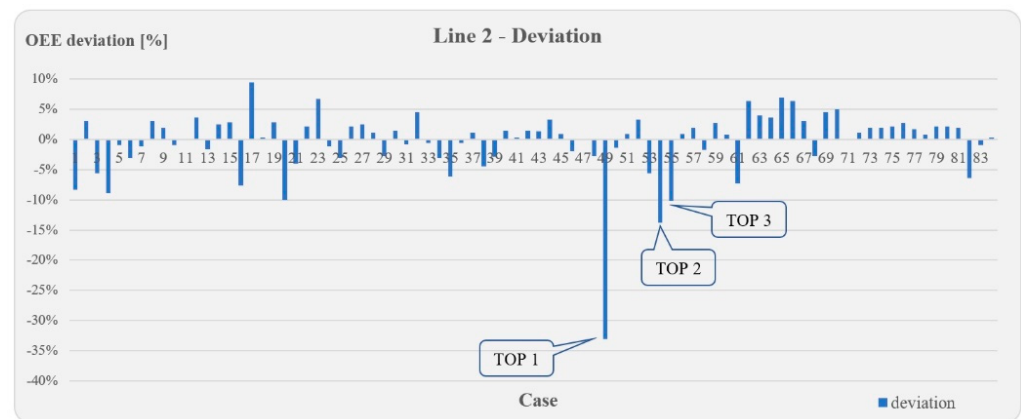


Figure 5. Deviation from the estimation for assembly line 2.

Table 5. TOP 3 deviation.

Assembly Line 1	Predicted	Achieved	Deviation	Reason
TOP 1 (case 42)	90.6%	30.0%	60.6%	Material missing
TOP 2 (case 27)	88.2%	49.1%	39.2%	Material missing
TOP 3 (case 77)	88.2%	59.4%	28.8%	Material missing
Assembly Line 2	Predicted	Achieved	Deviation	Reason
TOP 1 (case 49)	78.8%	45.8%	33.0%	Material missing
TOP 2 (case 54)	76.9%	63.1%	13.8%	Machine failure
TOP 3 (case 55)	73.1%	62.9%	10.2%	Machine failure

Root cause analysis found that the material missing was caused by lower inventory levels within the plant between production areas. Supermarkets in front of the assembly line are filled with materials manufactured by welding and painting machines where the minimum stock is inadequate. There are small deviations in the positive direction, which are due to, among other things, faster product changes and a lower number of short downtimes (<3 min) than planned. It is also important that exceeding the target OEE values is not an easy task, so that the planned breaks in the OEE calculation are included in the planned total time. A 30-min break, common in industrial practice, reduces the theoretically achievable 100% OEE by 6.25% ($30 \text{ min}/480 \text{ min} = 6.25\%$). Thus, the efficiency of the line can only be 93.75% in practice if no additional staff is employed during the planned breaks on the production line. The accuracy is characterized by the fact that 16% (line 1) and 20% (line 2) of the estimations were able to predict the OEE value with an accuracy of less than 1%.

Main events that make it impossible to accurately predict the OEE on the assembly line are:

- Machine failure and duration;
- Material missing and duration;
- Line stop due to raw material and/or subassembly quality;
- Line stop due to customer claim (root cause investigation);
- Urgent order and its consequence, or a program change;
- Other unforeseen events.

In the case where machine setters had to estimate the OEE value in advance for a week, none of them were able to predict this with an accuracy of less than two percent for any of the assembly line. In addition, in most cases, the estimation was not supported by a concept. This paper does not analyze the details of this case.

4.2. Takt Time Based Decision Trees

In the following, we will investigate how much more accurately OEE values can be predicted a few weeks in advance using a decision tree versus human estimation for a shift. The decision tree is based on the factors listed in Table 6.

Table 6. Factor description.

Abbr.	Factor	Short Description
RTT	Real takt time	Realized takt time of each product at the end of assembly line
PTT	Planned takt time	Planned takt time based on product type
PPT	Planned positive tolerance	Allowed takt time positive deviation
PNT	Planned negative tolerance	Allowed takt time negative deviation
MSDT	Maximum value of short downtime	Maximum duration of short downtime
T	Time	The time of prediction

From decision tree, each category can be defined in two steps. In the first step, the real takt time (RTT) and planned takt time (PTT) are compared, taking into account the planned negative (PNT) and positive tolerances (PPT) from which the following groups can be determined:

- Assembly operation with capacity extended assembly (temporarily more resources, e.g., multiple operators);
- Assembly operation with normal assembly process;
- Downtimes, which can be divided into two groups:
 - a. Short downtimes or micro stops;
 - b. Long downtimes.

The decision tree for classifying assembly operations and downtimes is shown in Figure 6.

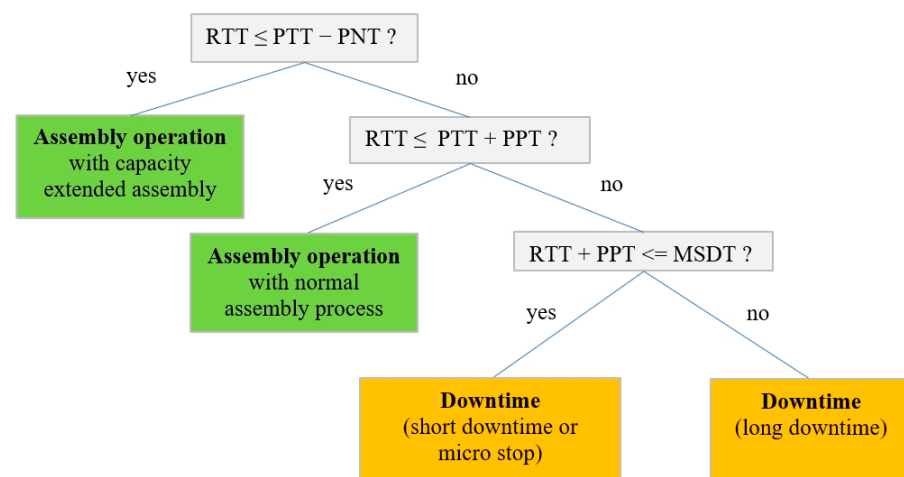


Figure 6. Decision tree for classification of assembly operation and downtime.

In the second step, long downtime is classified in terms of planned and unplanned downtimes. In addition to the individual takt time data of the products, the time stamp is also taken into account when defining each category. In other words, the planned time and date of shift change, and the duration and date of the two planned breaks for all 3 shifts, are taken into account. In the following, the further breakdown is based on whether a product change occurred during the downtime, because this also results in a separate group (planned downtime and type change). As a final step, unplanned downtimes are identified that include machine failures and material missing. In Figure 7, real industrial data were used to demonstrate the practical applicability of the decision tree.

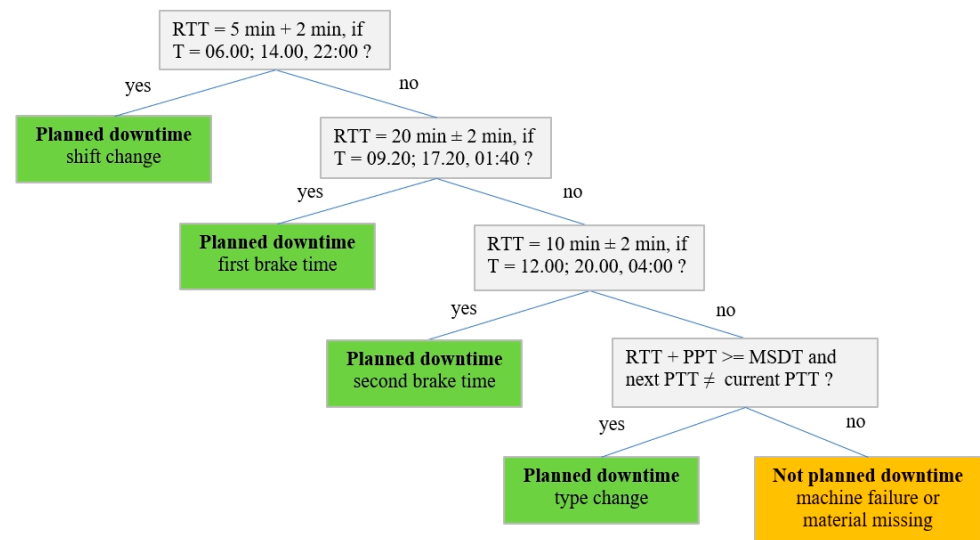


Figure 7. Decision tree for classification of different downtimes.

These two decision trees serve as the basis for the OEE prediction presented in the next section.

4.3. Target-Oriented OEE Prediction

The aim of the target-oriented OEE prediction model is to predict the OEE value of the next few weeks based on the decision trees presented earlier, taking into account the OEE values of the recent past with different weights, while also taking into account the planned OEE target. The target-oriented OEE prediction model is shown in Figure 8.

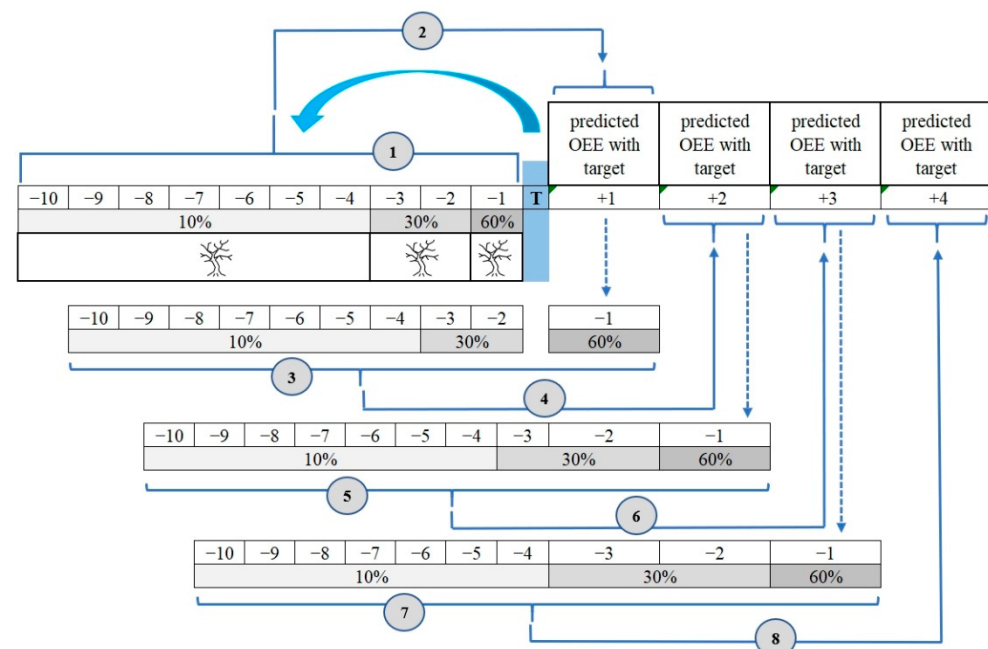


Figure 8. Target-oriented OEE prediction based on decision tree.

The model for each step is as follows:

1. At a given time (T), the last ten weeks are divided into three periods, as follows:
 - a. The OEE percentage for the last week (T − 1) is taken into account, with a weighting of 60%.

- b. The OEE percentage for the last second and third weeks ($T - 2$ and $T - 3$) is taken into account, with a weighting of 30%.
- c. From the last fourth week onwards, the OEE values for the period up to the last tenth week ($T - 4 \rightarrow T - 10$) are taken into account, with a weighting of 10%.

The groups described by the decision trees for each period and the values of their percentage distribution explain the magnitude and reasons for the OEE value.

2. The average of the OEE value obtained from the weighted values and the target set for the given assembly line will be the predicted OEE value for the first week ($T + 1$) after period T .
3. To determine the weekly OEE value for the period $T + 2$, the weighting ratio of 60%, 30% and 10% for ten weeks shall be modified as a rolling window, so that the 60% ratio is the value predicted in the section point 2 ($T + 1$) represents.
4. Based on section point 3, the average of the OEE value obtained from the weighted values and the target set for that assembly line will be the predicted OEE value for the second week ($T + 2$) after period T .
5. To determine the weekly OEE value for the period $T + 3$, the weighting ratio of 60%, 30% and 10% for ten weeks is modified like a rolling window, so that the 60% ratio is represented by the value predicted in section point 2 ($T + 1$), and the 30% rate will be the average of the $T - 1$ period and the $T + 1$ value predicted in section point 2.
6. Based on section point 5, the average of the OEE value obtained from the weighted values and the target set for that assembly line will be the predicted OEE value for the third week ($T + 3$) after period T .
7. To determine the weekly OEE value for the period $T + 4$, the weighting ratio of 60%, 30% and 10% for ten weeks is modified a rolling window, so that the 60% ratio is represented by the value predicted in section point 2 ($T + 1$), and the 30% ratio will be the average of the $T + 2$ predicted in section point 4 and the $T + 3$ predicted in section point 6.
8. Based on section point 7, the average of the OEE value obtained from the weighted values and the target set for that assembly line will be the predicted OEE value for the fourth week ($T + 4$) after period T .

The calculation method of the model shows the expected value of OEE, while the decision tree highlights why the value of OEE will be so high. The presented model is based on the fact that longer-term monitoring is not necessary, because in the current turbulent automotive market environment, the factors influencing the value of OEE may vary from week to week for the following reasons, among others:

- Material shortages (e.g., electronics chips, metal, plastic, etc.);
- Variable sales orders, variable quantity and product mix;
- Availability of resources (fluctuating energy prices, availability of labor, absence rate, etc.).

Besides these, other specific short- and long-term measures, e.g., previous technical and technological developments, the use of Lean methods and customer guidance also affect longer-term time series data, the impact of which is not always effective at present due to the decline in practical sustainability.

4.4. OEE Prediction Model in the Industrial Practice

For the hybrid assembly line 2 presented in Section 4.1 for the same period (week 43/2021–week 5/2022), the target-oriented OEE prediction model using the predicted values for the one-week forecast ($T + 1$) and the realized OEE values shows in Figure 9.

Based on the model, the four-week forecast descriptive statistics are illustrated in Table 7.

It is characteristic of the short-term prediction that the weekly values do not differ much for each parameter (e.g., minimum value, maximum value, mean, range and standard deviation). The authors publish more detailed classification distribution data based on decision trees in another scientific article [50].

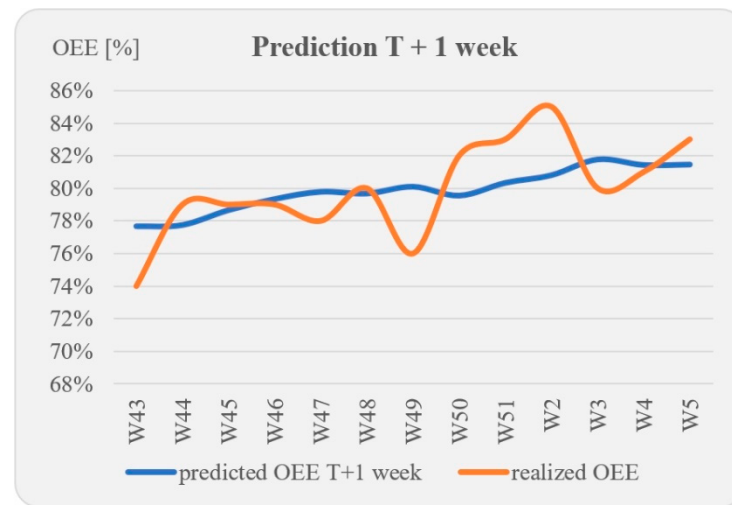


Figure 9. Predicted and realized OEE values.

Table 7. Descriptive statistics for target-oriented OEE prediction based on decision tree.

	T + 1 Week	T + 2 Week	T + 3 Week	T + 4 Week
Sample size	13	13	13	13
Minimum	77.7%	77.6%	77.7%	77.9%
Maximum	81.8%	81.7%	81.7%	81.5%
Mean	79.9%	79.8%	79.9%	79.9%
Range	4.1%	4.1%	4.0%	3.6%
Standard error	0.4%	0.4%	0.3%	0.3%
Standard deviation	1.3%	1.3%	1.2%	1.2%

5. Discussion

In this section, human prediction and target-oriented OEE prediction are compared. Human estimation is characterized by a wide forecast range, while target-oriented OEE prediction is a narrow range. This means that the maximum negative deviation of the forecast and the maximum positive deviation of the forecast compared to the realized one are much more accurate in the case of the target-oriented OEE prediction model. Both the accuracy of human prediction per shifts level and the accuracy of target-oriented OEE prediction for weekly levels are within 1%. The detailed comparison is shown in Table 8.

Table 8. Comparison of OEE estimation by human and target-oriented OEE prediction.

	Human Estimation	Target-Oriented OEE Prediction			
		T + 1	T + 2	T + 3	T + 4
Estimated period	1 shift	1 week	2 week	3 week	4 week
Range of prediction	38.5–80.8%	77.6–81.7%	77.6–81.7%	77.7–81.6%	77.9–81.5%
Average accuracy of prediction	0.4%	0.80%	0.76%	0.79%	0.79%
Maximum negative deviation	−33.0%	−4.1%	−3.6%	−3.6%	−3.5%
Maximum positive deviation	9.4%	4.1%	4.8%	5.5%	5.1%

The maximum negative and positive deviation is derived from the difference between the values of the estimated data and realized data in both the human prediction and target-oriented prediction model.

To sum up the advantages of the target-oriented OEE prediction model are as follows:

- Takes into account real and current circumstances;
- More accurate than human forecast with a smaller standard deviation (human prediction accuracy is more than two percent, target-oriented prediction less than one percent on a weekly basis);

- In contrast to most mathematical and statistical models, it ensures that the OEE cannot be greater than 100% (even for increasing trends, the calculation data is always less than 100%);
- Gives an accurate answer to the OEE value in terms of contributors (availability, performance and quality) with the help of decision trees; thus, in addition to forecasting, it also provides guidance on which groups to focus on in order to increase efficiency;
- Takes into account the goal to which the entire operational organization intends to go and allocates resources accordingly;
- Classification can be conducted using Excel; it does not require another tool or program (after MES or SQL query, the .csv file can be processed with Excel using various functions);
- It requires a small amount of human resources.

Compared to current research results, the presented method demonstrates the expected value of OEE, while the decision tree highlights why the value of OEE will be so high.

Based on these, the target-oriented OEE prediction model is excellent for estimating OEE percentages on semi-automatic assembly lines. The model can also be used to predict the efficiency of individual machines and full-automatic assembly lines in any industry. It could be a topic of further research to extend the target-oriented OEE prediction model for a longer estimation time, e.g., for two or three months. In addition, it is worth examining how the model behaves during different life cycles of the product, where efficiency is highly stochastic. Another way, could be in the future to develop the prediction model with a rule adaptation. During the prediction, recognized rules can be selected, confirmed and omitted to provide an even more accurate prediction.

6. Conclusions

This paper presented the estimation of Overall Equipment Effectiveness (OEE) at assembly lines by comparing human prediction with a takt time-based decision tree as one of the techniques of supervised machine learning. A detailed descriptive analysis was presented and a new prediction model was introduced and validated through a real industrial example. The aim of the target-oriented OEE prediction model is to predict the OEE value of the next few weeks based on the decision trees presented earlier, taking into account the OEE values of the recent past with different weights, while also taking into account the planned OEE target. The model can also be used to predict the efficiency of individual machines and full-automatic assembly lines in any industry. Using the model, the value of OEE can be predicted with an accuracy of within 1% on a weekly basis four weeks in advance.

Author Contributions: Conceptualization, P.D.; methodology, P.D. and J.J.; validation, P.D.; formal analysis, J.J.; investigation, P.D.; resources, P.D. and J.J.; data curation, P.D.; writing—original draft preparation, P.D.; writing—review and editing, P.D. and J.J.; visualization, P.D.; supervision, J.J.; funding acquisition, P.D. and J.J. All authors have read and agreed to the published version of the manuscript.

Funding: The APC was funded by Széchenyi István University, Győr, Hungary, Publication Support Program.

Institutional Review Board Statement: Not require ethical approval.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

- Enke, J.; Glass, R.; Kreß, A.; Hambach, J.; Tisch, M.; Metternich, J. Industrie 4.0 Competencies for modern production system. *Procedia Manuf.* **2018**, *23*, 267–272. [\[CrossRef\]](#)
- Kusiak, A. Smart Manufacturing. *Int. J. Prod. Res.* **2018**, *56*, 508–517. [\[CrossRef\]](#)
- Mrugalska, B.; Wyrwicka, M.K. Towards Lean Production in Industry 4.0. *Procedia Eng.* **2017**, *182*, 466–473. [\[CrossRef\]](#)
- Tao, F.; Qi, Q.; Liu, A.; Kusiak, A. Data-driven smart manufacturing. *J. Manuf. Syst.* **2018**, *48*, 157–169. [\[CrossRef\]](#)
- Mittal, S.; Khan, M.A.; Romero, D.; Wuest, T. Smart manufacturing: Characteristic, technologies and enables factors. *J. Eng. Manuf.* **2019**, *233*, 1342–1361. [\[CrossRef\]](#)
- Nagorny, K.; Lima-Monteiro, P.; Barata, J.; Colombo, A.W. Big data analysis in smart manufacturing: A review. *Int. J. Commun. Netw. Syst. Sci.* **2017**, *10*, 31–58. [\[CrossRef\]](#)
- 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. *J. Intell. Manuf.* **2020**, *31*, 1531–1558. [\[CrossRef\]](#)
- Almada-Lobo, F. The Industry 4.0 revolution and the future of Manufacturing Execution System (MES). *J. Innov. Manag.* **2016**, *3*, 16–21. [\[CrossRef\]](#)
- Mantravadi, S.; Moller, C. An overview of next-generation Manufacturing Execution System: How important is MES for Industry 4.0? *Procedia Manuf.* **2019**, *30*, 588–595. [\[CrossRef\]](#)
- Gyulai, D.; Kádár, B.; Monostori, L. Robust production planning and capacity control for flexible assembly lines. *IFAC-Pap. OnLine* **2015**, *48*, 2312–2317. [\[CrossRef\]](#)
- Gröger, C.; Niedermann, F.; Mitschang, B. Data mining-driven manufacturing process optimization. In Proceedings of the World Congress on Engineering, London, UK, 4–6 July 2012.
- Schuh, G.; Reinhart, G.; Prote, J.P.; Sauermann, F.; Horsthofer, J.; Oppolzer, F.; Knoll, D. Data mining definitions and applications for the management of production complexity. *Procedia CIRP* **2019**, *81*, 874–879. [\[CrossRef\]](#)
- Kusiak, A. Smart Manufacturing must embrace big data. *Nature* **2017**, *544*, 23–25. [\[CrossRef\]](#) [\[PubMed\]](#)
- Wuest, T.; Weimer, D.; Irgens, C.; Thoben, K.D. Machine learning in manufacturing: Advances, challenges, and applications. *Procedia Manuf. Res.* **2016**, *4*, 23–45. [\[CrossRef\]](#)
- Kang, Z.; Catal, C.; Tekinerdogan, B. Machine learning applications in production lines: A systematic literature review. *Comput. Ind. Eng.* **2020**, *149*, 106773. [\[CrossRef\]](#)
- Agard, B.; Kusiak, A. Data mining for subassembly selection. *J. Manuf. Sci. Eng.* **2004**, *126*, 627–631. [\[CrossRef\]](#)
- Backus, P.; Janakiram, M.; Mowzoon, S.; Runger, C.; Bhargava, A. Factory cycle-time prediction with a data-mining approach. *IEEE Trans. Semicond. Manuf.* **2006**, *19*, 252–258. [\[CrossRef\]](#)
- Harding, J.A.; Shahbaz, M.; Srinivas, S.; Kusiak, A. Data mining in manufacturing: A review. *J. Manuf. Sci. Eng.* **2006**, *128*, 969–976. [\[CrossRef\]](#)
- Meidan, Y.; Lerner, B.; Hassoun, M.; Rabinowitz, G. Data mining for cycle time key factor identification and prediction in semiconductor manufacturing. *IFAC Proc. Vol.* **2009**, *42*, 217–222. [\[CrossRef\]](#)
- Öztürk, A.; Kayaligil, N.E.; Özdemirel, N.E. Manufacturing lead time estimation using data mining. *Eur. J. Oper. Res.* **2006**, *173*, 683–700. [\[CrossRef\]](#)
- Espadinha-Cruz, P.; Godina, R.; Rodrigues, E.M.G. A review of data mining applications in semiconductor manufacturing. *Processes* **2021**, *9*, 305. [\[CrossRef\]](#)
- Lieber, D.; Stolpe, M.; Konrad, B.; Deuse, J.; Morik, K. Quality prediction in interlinked manufacturing processes based on supervised and unsupervised machine learning. *Procedia CIRP* **2013**, *7*, 193–198. [\[CrossRef\]](#)
- Tan, P.N.; Steinbach, M.; Kumar, V. *Introduction to Data Mining—Instructor's Solution Manual*; Pearson: London, UK, 2018.
- Witten, I.H.; Frank, E. *Data Mining—Practical Machine Learning Tools and Techniques*; Elsevier: San Francisco, CA, USA, 2005.
- Bhojani, H.S.; Bhatt, N. Data mining techniques and trends—A review. *Glob. J. Res. Anal.* **2016**, *5*, 252–254.
- Han, J.; Kamber, M.; Pei, J. *Data Mining—Concepts and Techniques*; Elsevier: Waltham, MA, USA, 2012.
- Qiao, X.; Jiao, H. Data mining techniques in analyzing process data: A didactic. *Front. Psychol.* **2018**, *9*, 2231. [\[CrossRef\]](#) [\[PubMed\]](#)
- Smola, A.; Vishwanathan, S.V.N. *Introduction to Machine Learning*; Cambridge University: Cambridge, UK, 2008.
- Choudhary, A.K.; Harding, J.A.; Tiwari, M.K. Data mining in manufacturing: A review based on the kind of knowledge. *J. Intell. Manuf.* **2009**, *20*, 501–521. [\[CrossRef\]](#)
- Nakajima, S. *Introduction to TPM: Total Productive Maintenance*; Productivity Press: Cambridge, UK, 1988.
- Corrales, L.C.; Lambán, M.P.; Korner, M.E.H.; Royo, J. Overall Equipment Effectiveness: Systematic literature review and overview of different approaches. *Appl. Sci.* **2020**, *10*, 6469. [\[CrossRef\]](#)
- Hedman, R.; Subramaniyan, M.; Almström, P. Analysis of critical factors for automatic measurement of OEE. *Procedia CIRP* **2016**, *57*, 128–133. [\[CrossRef\]](#)
- Sowmya, K.; Chetan, N. A review on effective utilization of resources using overall equipment effectiveness by reducing six big losses. *Int. J. Sci. Res. Sci. Eng. Technol.* **2016**, *2*, 556–562.
- Nguyen, M.N.; Do, N.H. Re-engineering assembly line with lean techniques. *Procedia CIRP* **2016**, *40*, 590–595. [\[CrossRef\]](#)
- Antoniolli, I.; Guariente, P.; Pereira, T.; Silva, F.J.G. Standardization and optimization of an automotive components production line. *Procedia Manuf.* **2017**, *13*, 1120–1127. [\[CrossRef\]](#)

36. Nguyen, T.L.; Le, M.T.; Vu, T.T.T.; Do, N.H. Lean line balancing for an electronics assembly line. *Procedia CIRP* **2016**, *40*, 437–442. [[CrossRef](#)]
37. Glass, R.; Seifermann, S.; Metternich, J. The spread of lean production in the assembly, process and machining industry. *Procedia CIRP* **2016**, *55*, 278–283. [[CrossRef](#)]
38. Dobra, P.; J3svai, J. Correlation between Overall Equipment Effectiveness (OEE) and Failure Mode and Effect Analysis (FMEA) at the semi-automatic assembly lines. In Proceedings of the IAC, Budapest, Hungary, 26–27 November 2021.
39. Heutmann, T.; Schmitt, R.H. Predictive control of a synchronized individual production. *Procedia CIRP* **2019**, *81*, 69–74. [[CrossRef](#)]
40. Lingitz, L.; Gallina, V.; Ansari, F.; Gyulai, D.; Pfeiffer, A.; Sihn, W.; Monostori, L. Lead time prediction using machine learning algorithm: A case study by a semiconductor manufacturer. *Procedia CIRP* **2018**, *72*, 1051–1056. [[CrossRef](#)]
41. Subramaniyan, M.; Skoogh, A.; Salomonsson, H.; Banglore, P.; Bokrantz, J. A data driven algorithm to predict throughput bottlenecks in a production system based on active periods of the machines. *Comput. Ind. Eng.* **2018**, *125*, 533–544. [[CrossRef](#)]
42. Anusha, C.H.; Umasankar, V. Performance prediction through OEE-model. *Int. J. Ind. Eng. Manag.* **2020**, *11*, 93–103. [[CrossRef](#)]
43. Schuh, G.; Potente, T.; Hauptvogel, A. Methodology for the evaluation of forecast reliability of production planning systems. *Procedia CIRP* **2014**, *17*, 469–474. [[CrossRef](#)]
44. Brunelli, L.; Masiero, C.; Tosato, D.; Beghi, A.; Susto, G.A. Deep learning-based production forecasting in manufacturing: A packaging equipment case study. *Procedia Manuf.* **2019**, *38*, 248–255. [[CrossRef](#)]
45. El Hassani, I.; El Mazgualdi, C.; Masrour, T. Artificial intelligence and machine learning to predict and improve efficiency in manufacturing industry. *arXiv* **2019**, arXiv:1901.02256.
46. Acosta, C.; Héctor, P.; Terán, C.; Arteaga, O.; Terán, M.B. Machine learning in intelligent manufacturing system for optimization of production cost and overall effectiveness of equipment in fabrication models. *J. Phys. Conf. Ser.* **2020**, *1432*, 012085. [[CrossRef](#)]
47. Felice, F.; Monfreda, S.; Petrillio, A.; Nenni, M.E.; Iannone, R.; Introna, V.; Giuiusa, A.; Carlo, F. *Operational Management*; The Open University of Hong Kong: Hong Kong, China, 2016.
48. Doerling, T.; Suresh, N.C. Forecasting and performance: Conceptualizing forecasting management competence as a high-order construct. *J. Supply Chain. Manag.* **2016**, *52*, 77–91. [[CrossRef](#)]
49. Lee, J.; Lapira, E.; Yang, S.; Kao, A. Predictive Manufacturing System—Trends of next-generation production systems. In Proceedings of the 11th IFAC Workshop on Intelligent Manufacturing System, Sao Paulo, Brazil, 22–24 May 2013.
50. Dobra, P.; J3svai, J. Predicting the impact of type changes on Overall Equipment Effectiveness (OEE) through machine learning. In Proceedings of the 1st IEEE International Conference on Internet of Digital Reality, Cuijk, The Netherlands, 23–24 June 2022.