Conference On Production Systems And Logistics
CPSL 2022

3rd Conference on Production Systems and Logistics

Approach To A Decision Support Method For Feature Engineering Of A Classification Of Hydraulic Directional Control Valve Tests

Christian Neunzig[1,2], Simon Fahle[1], Jürgen Schulz[2], Matthias Möller[2], Bernd Kuhlenkötter[1]
[1]*Chair of of Production Systems, Ruhr-University Bochum , Universitätsstraße 150 , 44801 Bochum, Germany*
[2]*Bosch Rexroth AG, Bexbacher Straße 72; 66424 Homburg, Germany*

Abstract

Advancing digitalization and high computing power are drivers for the progressive use of machine learning (ML) methods on manufacturing data. Using ML for predictive quality control of product characteristics contributes to preventing defects and streamlining future manufacturing processes. Challenging decisions must be made before implementing ML applications. Production environments are dynamic systems whose boundary conditions change continuously. Accordingly, it requires extensive feature engineering of the volatile database to guarantee high generalizability of the prediction model. Thus, all following sections of the ML pipeline can be optimized based on a cleaned database. Various ML methods such gradient boosting methods have achieved promising results in industrial hydraulic use cases so far. For every prediction model task, there is the challenge of making the right choice of which method is most appropriate and which hyperparameters achieve the best predictions. The goal of this work is to develop a method for selecting the best feature engineering methods and hyperparameter combination of a predictive model for a dataset with temporal variability that treats both as equivalent parameters and optimizes them simultaneously. The optimization is done via a workflow including a random search. By applying this method, a structured procedure for achieving significant leaps in performance metrics in the prediction of hydraulic test steps of directional valves is achieved.

Keywords

Predictive Quality; Machine Learning; Quality Control; Feature Engineering; Decision Support Method

1. Introduction

Rising computer capacities and increasing data availability are expanding the horizon for knowledge generation in production [1]. Politically and socially, increased requirements for sustainability and resource consumption are moving into the focus of companies [2]. More and more companies are perceiving digitization as an opportunity [3,4]. Advancing digitalization and high computing power are drivers for the progressive use of ML methods on production data. [5] One solution strategy for improving existing production systems is knowledge extraction from production data using ML [6]. Predictive quality describes the ability to make data-driven predictions of product- and process-related quality in the manufacture and use of physical products [7,8]. The use of ML for predictive quality control of product characteristics contributes to increasing the efficiency and sustainability of future manufacturing processes. [9] However, a success factor for a ML project is sufficient pre-sampling of production data concerning dynamically deviating boundary conditions during data generation and optimization of data quality. [10] The further sections of the ML pipeline can be optimized based on a cleaned database [11]. Various ML methods such as gradient boosting methods have so far achieved promising results in industrial hydraulic use cases [12].

For each prediction model task, there is the challenge of making the right choice of which algorithm with corresponding hyperparameters (HP) and which method of feature engineering (FE) is most appropriate to achieve the best predictions. FE is essential to the ML pipeline, as the overall performance of the model is highly dependent on the available features and enables the incorporation of domain knowledge into the ML pipeline [13]. By generating high-quality features, the effectiveness of an ML pipeline can be increased many times over an identical pipeline without dedicated FE and increases the interpretability of the trained model [14].

This work focuses on supervised learning, a subtype of ML in which the target values, so-called labels, of given data $\boldsymbol{\mathcal{D}} = \{(x_i, y_i) \in \mathbb{R}^d \times \mathbb{R}\}$ are known [15]. A ML model is fed with a training data set $\boldsymbol{\mathcal{D}}_{train}$ and evaluated against the validation data set $\boldsymbol{\mathcal{D}}_{valid}$, where x_i is respectively the input with dimension d and y_i is the target [16]. For the given dataset $\boldsymbol{\mathcal{D}}$, the goal is to find the algorithm $A_j \in \boldsymbol{\mathcal{A}}$ from hypothesis space $\mathcal{A}$ with HP combination $\lambda_j \in \Lambda$ from HP space Λ that achieves the highest accuracy under the premise of highest possible generalizability. The dataset $\boldsymbol{\mathcal{D}}$ is only an extract from the population, thus the model attempts to achieve the most accurate representation of the output for a given input by approximating a transfer function [16,17].

In this paper, a decision-support method is developed for selecting appropriate FE techniques while choosing the best HP combination of a predictive model for a dataset with temporal variability and validated on a real industrial value stream of directional control valves at Bosch Rexroth. By applying this method, structured guidance is given and the automation of manual efforts for achieving significant leaps in the performance metrics of an extreme gradient boosting (XGB) classifier is achieved. The method provides a qualitative evaluation of different FE techniques in hydraulic directional valve manufacturing and achieves significantly more accurate predictions than sequential pipeline optimization as shown in the experiment section.

2. Related Work

The ML pipeline and the simultaneous selection of the prediction algorithm and HP are explained. The importance of Fs application in various hydraulic applications will be addressed.

2.1 Combined Algorithm Selection and HP optimization problem

Within the ML pipeline, the task blocks of data preparation, the FE, model generation and evaluation are run through sequentially [18]. HÜTTER ET AL. describe the problem of automatic optimization of ML pipelines, where the choice is made between the correct ML algorithms and corresponding categorical HP, as the combined algorithm selection and HP optimization problem (CASH), see equation (1). The loss function is minimized against all folds k of the cross-validation (CV) for the training data set $\boldsymbol{\mathcal{D}}_{train}$ with $\boldsymbol{\mathcal{D}}_{train}^{(i)} = \boldsymbol{\mathcal{D}}/\boldsymbol{\mathcal{D}}_{valid}^{(i)}$ from all models considered to limit the effect of overfitting and increasing the generalizability [11].

$$\boldsymbol{A}_{\lambda^*}^{*} \in \underset{A_j \in \mathcal{A},\ \lambda_j \in \Lambda}{\operatorname{argmin}} \frac{1}{k} \sum_{i=1}^{k} \mathcal{L}(\boldsymbol{A}_{\lambda}^{(j)}, \boldsymbol{\mathcal{D}}_{train}^{(i)}, \boldsymbol{\mathcal{D}}_{valid}^{(i)}) \tag{1}$$

ZÖLLER ET AL. extend the CASH problem to optimize the pipeline $\mathrm{P} \in \mathcal{P}$ with the pipeline structure $g \in \mathcal{G}$ from a set of valid pipeline structures $\mathcal{G}$, where $|g|$ is the length of the pipeline, see (2) [14].

$$(g, \boldsymbol{A}, \boldsymbol{\lambda})^* \in \underset{A_j \in \mathcal{A}^{|g|},\ \lambda_j \in \Lambda}{\operatorname{argmin}} \frac{1}{k} \sum_{i=1}^{k} \mathcal{L}(P_{g, \boldsymbol{A}, \boldsymbol{\lambda}, \boldsymbol{\mathcal{D}}_{train}^{(j)}}, \boldsymbol{\mathcal{D}}_{valid}^{(i)}) \tag{2}$$

2.2 Feature Engineering & CASH Optimization Methods for Hydraulic Use Cases

FE is about generating and selecting features from a given data set for the subsequent modelling step. FE can be split into three sub-tasks: feature extraction (FEX), feature construction (FC) and feature selection (FS) [19]. FS defines a subset of the feature set to speed up subsequent ML model training and improve its performance by removing redundant or misleading features. Simple domain-agnostic filtering approaches for FS are based on information theory and statistics. Methods such as univariate selection, variance threshold, feature importance, correlation matrices, or stability selection are integrated components of modern automated ML frameworks and are selected using standard CASH methods such as XGB algorithms [14]. Effective optimization approaches use random search or test the performance set exhaustively in a grid search also known as full factor design [16,20]. Advantages of random over grid search include easier parallelization and flexible resource allocation, since aborted iterations do not cause data holes and the combinations can be calculated independently [11]. Moreover, commonly used optimization methods contain reinforcement learning, evolution-based algorithm, and gradient descent, surrogate model-based optimization [18,21].

FS are using wrapper functions searches for the best feature subset by testing its performance on a given ML algorithm [19]. In heuristic approaches, individual features are added iteratively. Both forward and backward selection, as well as a combination of both, can be performed to select a subset of features. [18] In embedded methods, FS is directly integrated into the training process of an ML model. Many ML models, such as the Random Forest, provide some form of feature ranking that can be used. Similarly, embedded methods can be used in combination with FEX and FC [14]. Another optimization alternative is to use genetic programming in combination with prediction algorithms to identify a functioning feature subset [22]. TRAN ET AL. also used genetic programming to artificially construct novel features. In addition, information about how many times each feature was used during FC is reused to obtain feature importance [23]. KATZ ET AL. propose to compute meta-features for each novel feature, such as the diversity of values or the mutual information with the other features. Using a pre-trained classifier, the influence of a single feature can be predicted to select only promising features [24].

The prediction of internal leakage using discrete quality data has not yet been considered in science. The following are similar hydraulic industrial use cases whose approach has been considered in this work. LEI ET AL. recommend the combined use of principal component analysis (PCA) for dimensionality reduction of timeseries data with the classification and regression trees, random forests, and XGB, respectively, for the fault diagnosis model of the hydraulic valve [25]. KLUSCH ET AL. propose various statistical signal preprocessing steps such as correlation analysis and exploit this statistical characteristic in the form of features in a linear discriminant analysis and a k-nearest neighbor classifier for timeseries fault prediction [26]. HELWIG ET AL. identify times of equal boundary conditions in the prediction of hydraulic leakage flows and consider them during cross-validation and vary the weighting of the data within different time intervals in timeseries data [27,28].

3. Use Case: Internal Leakage Flow of Directional Control Valves

In this industrial use case at Bosch Rexroth, quality prediction of customer characteristics of directional control valves manufactured in Homburg is realized based on geometric gauge blocks from machining, mating data from assembly and hydraulic sensor data from end-of-line testing, shown in Figure 1. The target variable of this use case is the physically unavoidable internal leakage volume flow between spool and housing bore of a directional control valve [29]. The objective is subject to some uncertainty due to non-measurable variables and manufacturing tolerances. Previous work has attempted to understand the variables affecting leakage primarily through CFD simulations [30].

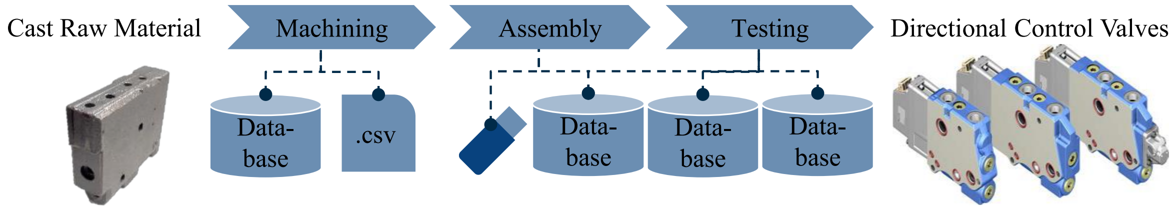

Figure 1: Value Chain of "Directional Control Valve" in Homburg.

An increased leakage volume flow leads to an unintentionally faster lowering of the load and involves dangers for people and property, see Figure 2a. The upper limit of leakage is secured in one of over 60 test steps as a safety-critical product characteristic within the scope of the final hydraulic inspection. In Figure 2b the 2-D view of the relevant valve area of the directional control valve is displayed. For the directional valve to function, the spool with different shoulders must be able to move within the housing bore, so that a ring gap between the bore and the spool, which is smaller than 20 micrometres, is inherent in the system, see Figure 2b [33]. The leakage measurement is carried out indirectly in the form of a pressure drop measurement at the pressure chamber of customer port A, which is reduced due to the leakage volume flow through the ring gap into the tank.

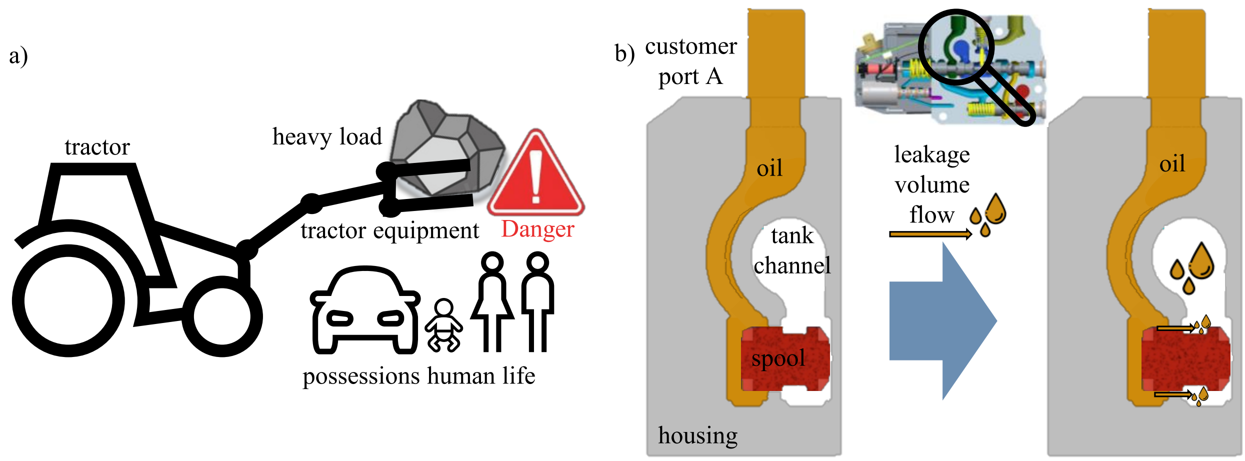

Figure 2: (a) safety-critical customer characteristic; (b) leakage volume flow as a pressure drop measurement.

4. Proposed Methodology & Experiments

The decision support method for selecting appropriate FE techniques is presented, applied to the hydraulic use case and finally the output is compared to a default as well as a tuned baseline model.

4.1 Decision Support Method for Selection of FE techniques & Pipeline Optimization

The proposed method for optimizing the ML pipeline and selecting the best fitting FE techniques builds on the CASH problem description of ZÖLLER ET AL [14]. The model matrix $\mathcal{M}$ forms the hypothesis space for the model candidates solving a prediction problem and consists of the three vectors each one for FE space $\mathcal{E}$, algorithm space $\mathcal{A}$ and HP space Λ, see Figure 3 and (4). The HP selection λ_w and partially the FE selection e_u is determined by the selection of the algorithm A_v since some algorithms will not run without appropriate FE. Thus, different types of algorithms require different numbers of FEs, so the matrix is the column dimension from the maximum of u, v, and w with u, v, w $\in \{1, 2, \ldots, n\}$. The matrix is used to produce different combinations from the three vectors. Empty matrix cells contain a zero and are not considered as a combination.

$$\text{Model } \mathbf{M} \in \boldsymbol{\mathcal{M}}_{3\,\mathrm{x}\,n} = \begin{pmatrix} e_{u\,\in\,\{1,2,\ldots,n\}} \in \boldsymbol{\mathcal{E}} \\ A_{v\,\in\,\{1,2,\ldots,n\}} \in \boldsymbol{\mathcal{A}} \\ \lambda_{w\,\in\,\{1,2,\ldots,n\}} \in \Lambda \end{pmatrix} = \begin{pmatrix} e_1 & e_2 & \ldots & e_n \\ A_1 & A_2 & \ldots & A_n \\ \lambda_1 & \lambda_2 & \ldots & \lambda_n \end{pmatrix} \begin{array}{l} \text{FE Space } \boldsymbol{\mathcal{E}} \\ \text{Algorithm Space } \boldsymbol{\mathcal{A}} \\ \text{HP Space } \Lambda \end{array}$$

Figure 3: Model Matrix of the Proposed Method for FE & Model Optimization.

The workflow of the method provides guidance on the structured approach to pipeline optimization and a qualitative trade-off of different FE techniques, see Figure 4.

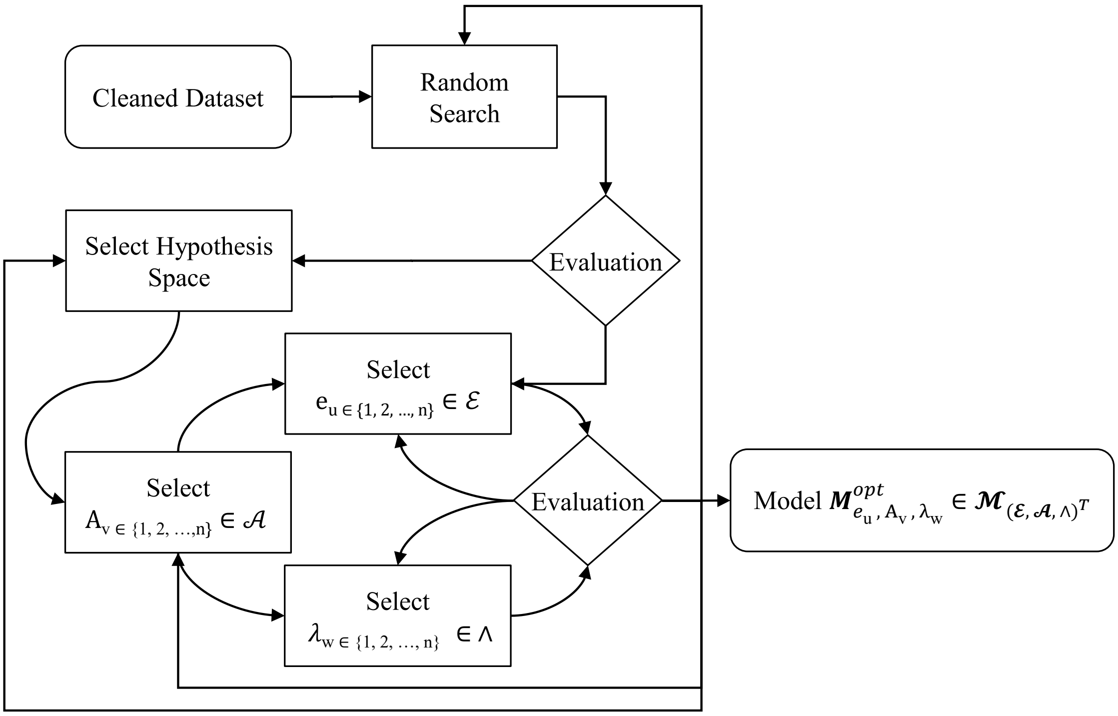

Figure 4: Flow Chart of the Proposed Method for FE & Model Optimization.

Various iteration loops illustrate the high combinability of the parameters. The random search is optimized with the given number of iterations. Based on a cleaned database, a random search is used to treat the FE techniques and the HP as equivalent optimization parameters. For the FE technique, this is partly achieved by providing a grid with Boolean data type. The inherent parameters of the FE techniques are treated as HP of the algorithms in the hypothesis space and are optimized according to the given grid. The procedure allows CV with k folds with different size as well for all combinations $M \in \mathcal{M}$. For optimization, the loss function from equation (4) is minimized against the validation data $\boldsymbol{\mathcal{D}}^{(i)}_{valid}$.

$$\boldsymbol{M}^{opt}_{e_u, A_v, \lambda_w} \in \boldsymbol{\mathcal{M}}_{(\mathcal{E},\mathcal{A},\Lambda)^T} = \boldsymbol{M}^{opt}_{e_u, A_v, \lambda_w} \in \underset{\boldsymbol{M}^{u,v,w}_{e_u \in \mathcal{E}, A_v \in \mathcal{A}, \lambda_w \in \Lambda}}{\operatorname{argmin}} \frac{1}{k}\sum_{i=1}^{k} \mathcal{L}(\boldsymbol{M}^{(u,v,w)}_{\mathcal{E}_u, \mathcal{A}_v, \Lambda_w}, \boldsymbol{\mathcal{D}}^{(i)}_{train}, \boldsymbol{\mathcal{D}}^{(i)}_{valid}) \quad (4)$$

4.2 FE Space: Derivation of used FE Techniques

Two core problems exist in the present production dataset: high dimensionality and high volatility over time. The production dataset consists of 11,652 data series with 1,052 features, which is a relatively unfavourable row-column ratio. The input features consist of geometric characteristics from machining, pairing information from assembly and sensor information from previous inspection steps. For this problem in a volatile high dimensional production data set, the XBG is a reasonable candidate of an algorithm [34]. The first task to be solved is the dimensionality reduction, see Figure 7. Preliminary studies have shown that

PCA is far superior to linear discriminant analysis just as forward selection is much better suited than backward selection for this data set. The final test data is composed as a matrix of the test steps (row-wise called PCA-A) and the sensors (column-wise called PCA-B). Second, the challenge of the changing boundary conditions of the production system manifest itself in value jumps in the data, as illustrated in Figure 5 for the scatter plots of two features. It is critical to recognize that the data jumps occur at different times for different feature columns. It is necessary to prevent the creation of new models each time the boundary condition changes, otherwise the amount of available data decreases drastically. Therefore, the change must be incorporated into the model by adjusting the same-sized fold CV. In future studies, identification with change point detection will be explored more intensively.

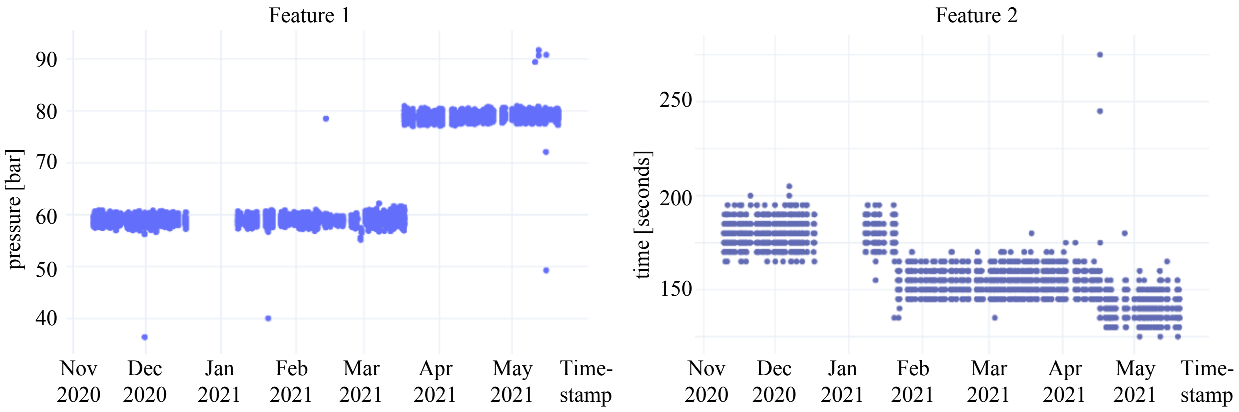

Figure 5: Exemplary scatter plots of two features with significant value jumps.

The 11 intervals with the same boundary conditions are indicated with the number of instances per interval, as shown in Figure 6. In this work, the interval-dependant timeseries 10-fold CV for each interval can be included in the presented workflow in Figure 4.

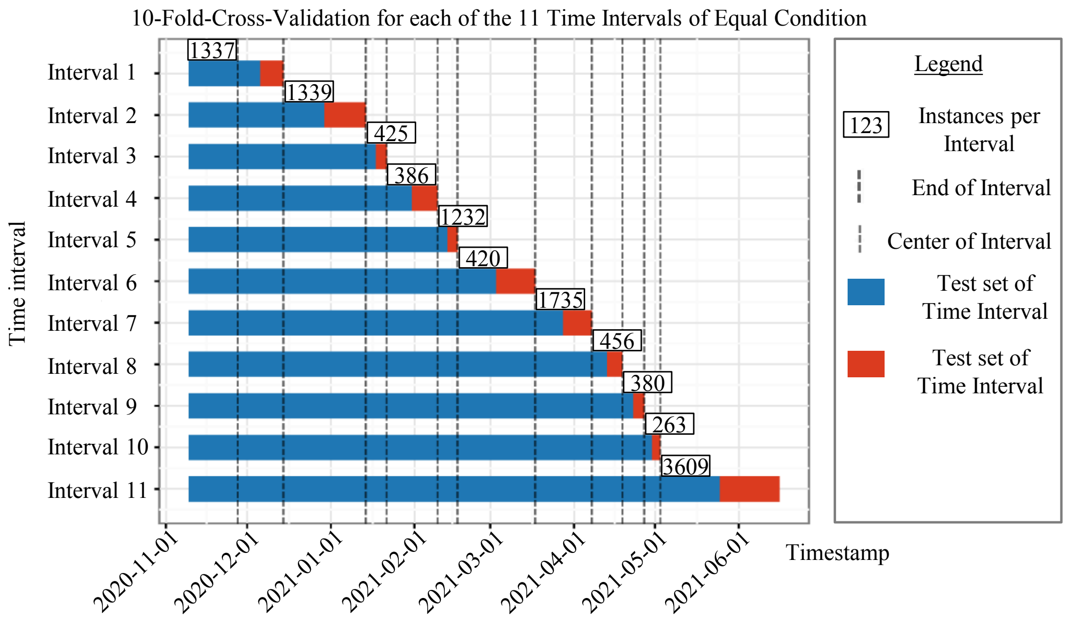

Figure 6: Timeseries-10-Fold-CV for folds for each time interval of equal systemic boundary conditions.

For the first interval, an 80/20 train-test split is performed. The following splits contain the previous splits and additionally the first half of the current split. The second half of the considered split is used as test data set. Within each interval, a 10-fold CV is performed and $\mathcal{D}_{valid}^{(i)}$ is created. To account for the temporal variability of the data set, further weighting and consideration of different subsets are also applied using a window-based approach. More recent data points are weighted stronger within the loss function, applying exponential smoothing familiar from time series analysis and devaluing data points from earlier periods by a factor, using $\beta = (1-\beta)^k$ with $\beta \in \{0,1\}$ and k = 0 for the current interval and k = 1 for the following and so on, see Figure 7d. For the window-based approach, only the number n most recent data series backward to the prediction time are considered in the model to predict the state at the current prediction time, see Figure 7b. In addition, the data per interval are each transformed with centering and scaling, see Figure 7c.

All in all, in the proposed method a forward selection, a PCA depending on each test step (PCA-A) or sensor (PCA-B) are given as input for a dimensionality reduction and further compared to no application of dimensionality reduction at all. To evaluate the predictive power of data in the near and distant past of the prediction state, different amounts of data are utilized as input and a pre-optimized weighting function is applied. This weighting function is randomly activated and deactivated in the method, see Figure 7d. In addition, the data for all intervals is transformed to a comparable level via scaling and centering, see Figure 7d.

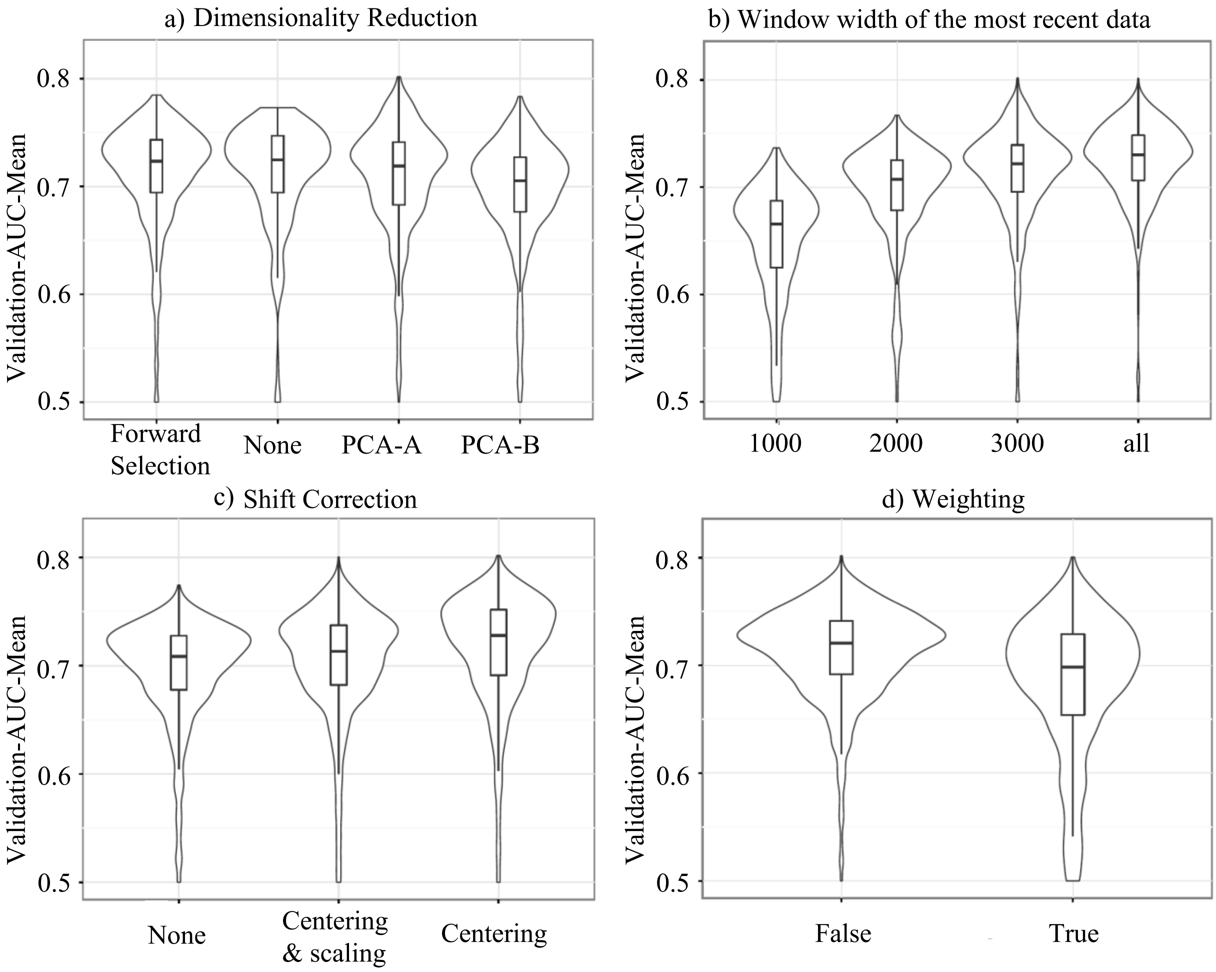

Figure 7: Qualitative Comparison of the FE techniques by Violin Plots.

4.3 Application and Evaluation of the Decision Support Method

A major strength of the presented method is the possibility of comparing several FE techniques while simultaneously applying and simultaneously optimizing the HP. Another advantage is the parallel calculation of the iterations within the method since the iterations are calculated independently. A suitable plot to compare the FE methods is the violin plot including frequencies and boxplot of AUC for different FE techniques, see Figure 7.

Thereby it turns out that the PCA-A with an AUC of 80 % produces the best candidate for the dimensionality reduction technique. Centering is the most effective technique for correcting the shift for this data set, as the best model shows the maximum AUC value of 80 % and averages 3 percentage points better than the variant with scaling. The weighting function should preferably be disabled in combination with listed FE techniques. Interesting to mention that the consideration of the last 3000 data series from the prediction time is not so much worse than the combination with consideration of the complete data set.

The proposed decision support method for FE techniques is applied to the hydraulic use case and compared with two baseline approaches to evaluate the suitability of the method, see Table 1. Many binary classification algorithms calculate the classification boundary and classify according to whether the value is above a certain threshold or not. The AUC is the trade-off between true-positive rate and false-positive rate that applies to all possible thresholds, not just the threshold chosen by the modelling technique. Different classification goals may result in one point on the curve being better suited for one task than another for different task. Hence, looking at the AUC is one way to evaluate the model regardless of the choice of a threshold. The F1-score brings precision and recall harmonized and is therefore very suitable as an evaluation of unbalanced data sets. The first base model (A) was applied to the data set with standard HP and without FE techniques. The second base model (B) consists of matched HP and preselected FE techniques, each of which produced the best results when applied individually. The application of centering, a weighting function with beta $\beta = 0.8$ and a PCA-B is performed for all data in model (B). The optimized model $\boldsymbol{M}^{opt}_{e_u, A_v, \lambda_w}$ is the best candidate of the proposed method after 5000 iterations with described FE input in 4.2, tuned HP of an XGB classifier and classifier itself. Sequential pipeline optimization and combination of preselected methods show a large effect on F1-score and AUC for unseen data. However, the application of the proposed method shows a significant jump in terms of F1-score by 10.4 percentage points and on the AUC of 8.3 percentage points for the unseen data compared to the sequential optimization.

Table 1: Comparison of different models for validation of the proposed method.

Model type	**FE**	**HP**	**AUC Train**	**AUC Test**	**F1-Score Train**	**F1-Score Test**
baseline model (A)	no methods	default	60,6 %	55,8 %	6.28 %	**3.33 %**
baseline model with FE (B)	pre-choice	tuned	68,7 %	**71,9 %**	11.3 %	**12.96 %**
optimized model $\boldsymbol{M}^{opt}_{e_u, A_v, \lambda_w}$	tuned	tuned	99,9 %	**80,2 %**	24.78 %	**23.36 %**

5. Conclusions and Future Work

The proposed decision support method produces the best combination of FE techniques and tuned HP for the XGB classifier on the time volatile production data of the industrial use case for predicting the leakage volume flow of directional control valves and outperforms the sequential optimization by about 10 percentage points in F1-score. From the qualitative review of various combinations of FE techniques, the combination of PCA-A, a data centring and a maximum window width with deactivated weighting emerged as the best FE combination.

Future work will validate a generalized suitability of the decision support method by applying it to additional algorithms and data sets. In addition, the optimization of the developed model and its implementation in production will be investigated. Feature construction approaches will be pursued to improve the usefulness of the model for series implementation by incorporating more predictive information into designed features.

References

[1] Posoldova, A., 2020. Machine Learning Pipelines: From Research to Production. IEEE Potentials 39 (6), pp. 38–42.

[2] Trauth, D., Bergs, T., Prinz, W., 2021. Monetarisierung von technischen Daten. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 1–734.

[3] Engels, B., Goecke, H., 2019. Big Data in Wirtschaft und Wissenschaft: Eine Bestandsaufnahme: IW-Analysen, No. 130. ECONSTOR, pp. 1–62.

[4] Lorenz, R., Benninghaus, C., Friedli, T., Netland, T.H., 2020. Digitization of manufacturing: the role of external search. IJOPM 40 (7/8), pp. 1129–1152.

[5] Krauß, J., Frye, M., Beck, G.T.D., Schmitt, R.H., 2019. Selection and Application of Machine Learning- Algorithms in Production Quality, in: Beyerer, J., Kühnert, C., Niggemann, O. (Eds.), Machine Learning for Cyber Physical Systems, vol. 9. Springer Berlin Heidelberg, pp. 46–57.

[6] Tao, F., Qi, Q., Liu, A., Kusiak, A., 2018. Data-driven smart manufacturing. Journal of Manufacturing Systems 48, pp. 157–169.

[7] Bergs, T., 2020. Internet of Production - Turning Data into Value, pp. 229–239.

[8] Wang, J., Ma, Y., Zhang, L., Gao, R.X., Wu, D., 2018. Deep learning for smart manufacturing: Methods and applications. Journal of Manufacturing Systems 48, pp. 144–156.

[9] Krauß, J., Pacheco, B.M., Zang, H.M., Schmitt, R.H., 2020. Automated machine learning for predictive quality in production. Procedia CIRP 93, pp. 443–448.

[10] Klinkenberg, R., 2005. Meta-Learning, Model Selection, and Example Selection in Machine Learning Domains with Concept Drift. Lernen, Wissensentdeckung und Adaptivität (LWA), pp. 1–8.

[11] Hutter, F., Kotthoff, L., Vanschoren, J., 2019. Automated Machine Learning. Springer International Publishing, Cham, pp. 84–95.

[12] Weichert, D., Link, P., Stoll, A., Rüping, S., Ihlenfeldt, S., Wrobel, S., 2019. A review of machine learning for the optimization of production processes. International Journal Advanced Manufacturing Technology 104 (5-8), pp. 1889–1902.

[13] Chu, X., Ilyas, I.F., Krishnan, S., Wang, J., 2016. Data Cleaning, in: Özcan, Koutrika et al. (Hg.)– Proceedings of the 2016 International, pp. 2201–2206.

[14] Zöller, M.-A., Huber, M.F., 2021. Benchmark and Survey of Automated Machine Learning Frameworks. jair, pp. 409–472.

[15] Hastie, Trevor, Jerome Friedman und Robert Tibshirani, 2001. The Elements of Statistical Learning.: Data Mining, Inference, and Prediction, pp. 29–39.

[16] Feurer, M., Hutter, F., 2019. Automated Machine Learning: Hyperparameter Optimization. Springer International Publishing, Cham, pp. 7–33.

[17] Raschka, S., 2018. Model Evaluation, Model Selection, and Algorithm Selection in Machine Learning, pp. 1–49.

[18] He, X., Zhao, K., Chu, X., 2021. AutoML: A survey of the state-of-the-art. Knowledge-Based Systems 212, pp. 1–37.

[19] Motoda, H., Liu, H., 2002. Feature Selection Extraction and Construction. Communication of IICM (Institute of Information and Computing Machinery, Taiwan) Vol 5., pp. 67–72.

[20] Montgomery, D.C., 2013. Design and analysis of experiments, John Wiley & Sons. pp. 1–757.

[21] Nicolo Fusi, Rishit Sheth, Melih Elibol, 2018. Probabilistic Matrix Factorization for Automated Machine Learning. Neural Information Processing Systems, Montréal, Canada.

[22] Olson, R.S., Moore, J.H., 2019. TPOT: A Tree-Based Pipeline Optimization Tool for Automating Machine Learning, in: Hutter, F., Kotthoff, L., Vanschoren, J. (Eds.), Automated Machine Learning. Springer International Publishing, Cham, pp. 151–160.

[23] Tran, B., Xue, B., Zhang, M., 2016. Genetic programming for feature construction and selection in classification on high-dimensional data. Memetic Comp. 8 (1), pp. 3–15.

[24] Katz, G., Shin, E.C.R., Song, D., 2016. ExploreKit: Automatic Feature Generation and Selection, in: 2016 IEEE 16th International Conference on Data Mining (ICDM). 2016 IEEE 16th International Conference on Data Mining (ICDM), Barcelona, Spain. 2016. pp. 979–984.
[25] Lei, Y., Jiang, W., Jiang, A., Zhu, Y., Niu, H., Zhang, S., 2019. Fault Diagnosis Method for Hydraulic Directional Valves Integrating PCA and XGBoost. Processes 7 (9), pp. 589 –607.
[26] Klusch, M., Meshram, A., Schuetze, A., Helwig, N., 2015. iCM-Hydraulic, in: Proceedings of the 11th International Conference on Semantic Systems. SEMANTiCS '15: 11th International Conference on Semantic Systems, Vienna Austria. 2015. ACM, New York, NY, USA, pp. 81–88.
[27] Helwig, N., 2018. Zustandsbewertung industrieller Maschinen mittels multivariater Sensordatenanalyse. Dissertation, pp. 1–210.
[28] Helwig, N., Pignanelli, E., Schutze, A. Condition monitoring of a complex hydraulic system using multivariate statistics, in: 2015 IEEE International Instrumentation, pp. 210–215.
[29] Matthies, H.J., Renius, K.T., 2014. Einführung in die Ölhydraulik. Springer Fachmedien Wiesbaden, pp. 38–48.
[30] Posa, A., Oresta, P., Lippolis, A., 2013. Analysis of a directional hydraulic valve by a Direct Numerical Simulation using an immersed-boundary method. Energy Conversion and Management 65, pp. 497–506.
[31] Morris, N., Rahmani, R., Rahnejat, H., King, P.D., Fitzsimons, B., 2013. Tribology of piston compression ring conjunction under transient thermal mixed regime of lubrication. Tribology International 59, pp. 248–258.
[32] Neunzig, C., Fahle, S., Kuhlenkötter, B., Möller, M., 2021. Feature Engineering For A Cross-process Quality Prediction Of An End-of-line Hydraulic Leakage Test Using An Experiment Sample, 2nd Conference on Production Systems and Logistics, pp. 156–166.
[33] Amirante, R., Del Vescovo, G., Lippolis, A., 2006. Flow forces analysis of an open center hydraulic directional control valve sliding spool. Energy Conversion and Management 47 (1), pp. 114–131.
[34] Mangal, A., Kumar, N., 2016. Using big data to enhance the bosch production line performance: A Kaggle challenge., in: 2016 IEEE International Conference, pp. 2029–2035.

Biography

Christian Neunzig has been a member in the Bosch PhD program since 2019. Previously, he completed a master's degree in production engineering and a bachelor's degree in mechanical engineering at RWTH Aachen University. His research focus is applied machine learning on directional control valve manufacturing data.

Simon Fahle has been a member of the Chair of Production Systems (LPS) at Ruhr University Bochum since 2019. He completed a bachelor's and master's degree in mechanical engineering at Ruhr-University Bochum. His primary research topics are machine learning, time series and radial-axial ring rolling.

Jürgen Schulz is manager of the process planning group and assembly and testing technology at the Bosch Rexroth plant in Homburg. Mr. Schulz received diploma (Dipl.-Ing.) from the University of Applied Sciences in Saarbrucken (HTW).

Matthias Möller is head of the Technical Functions department at the Bosch Rexroth plant in Homburg. Mr. Möller received his doctorate (Dr.-Ing.) and diploma (Dipl.-Ing.) from the Technical University of Dortmund.

Bernd Kuhlenkötter was responsible for product management and technology at ABB Robotics Germany until 2009. In 2009, Bernd Kuhlenkötter took over the professorship for "Industrial Robotics and Production Automation" at the Technical University of Dortmund. Univ. Prof. Dr.-Ing. Bernd Kuhlenkötter has held the professorship for "Production Systems" at the Ruhr University Bochum since 2015.