International Conference on Applied Energy 2019
Aug 12-15, 2019, Västerås, Sweden
Paper ID: 589

GENERIC MACHINE LEARNING APPROACH FOR VERY SHORT TERM LOAD FORECASTING OF PRODUCTION MACHINES

Jessica Walther[1*], Bastian Dietrich[1], Eberhard Abele[1]

1 Institute of Production Management, Technology and Machine Tools (PTW), Technische Universität Darmstadt, Otto-Berndt-Str. 2, 64287 Darmstadt, Germany

* Corresponding author. Tel.: +49 (0) 6151 16 20859. E-mail address: j.walther@ptw.tu-darmstadt.de

ABSTRACT

With the ongoing digitalization of industrial production, an increasing number of energy measuring points are installed in manufacturing environments which enable promising use cases for machine learning applications. This paper presents a generic machine learning approach to forecast the very short term load of production machines which can be utilized as decision support basis for control schemes and measures to decrease energy costs. The presented approach is developed and evaluated on production machines of the ETA research factory at the Technische Universität Darmstadt. The results indicate that the developed approach is feasible and creates a precise very short term load forecasting model for different production machines.

Keywords: Load forecasting, machine learning, feature selection, feature engineering, artificial neural networks

1. INTRODUCTION AND MOTIVATION

The industrial sector is the largest energy consumer worldwide. Focusing on electricity the industry accounts for 41.6% of the total consumption in 2016 [1]. In terms of electric energy consumption on a national level, the German industry sector alone consumes over 42.6% of total country's consumption. The manufacturing sector represents a subset of the industrial sector and is responsible for 93.7% of the industrial sectors energy consumption [2]. According to [3] these high levels of consumed energy during manufacturing lead to an economic motivation for companies to increasing their energy efficiency.

NONMENCLATURE

Abbreviations	
ANN	Artificial neural network
FC	Feature Construction
FE	Feature Engineering
FS	Feature Selection
MAD	Median Absolut Deviation of the Median
PLC	programmable logic controller
RF	Random Forest
VIF	Variance Inflation Factor

Against this background an increasing number of energy measuring points are installed in manufacturing environments, generating vast amounts of data [4]. Furthermore, modern production machines provide a large amount of process data. However, comprehensive data analysis is only partially implemented in the manufacturing industry. In particular, machine-internal data such as programmable logic controller (PLC) and bus data are usually not included in the analysis.

These energy data sources represent a promising machine learning application to extract valuable information. Load forecasting represents a use case of special interest in the manufacturing industry, as the information of the future energy demand is required for various subsequent optimization measures to decrease energy costs e.g. by peak load management, optimal energy purchasing or demand response applications [5,6]. Against the background of increasing data availability, machine learning based load forecasts are a promising solution to exploit theses potentials.

Selection and peer-review under responsibility of the scientific committee of the 11th Int. Conf. on Applied Energy (ICAE2019).
Copyright © 2019 ICAE

On that account, a machine learning approach to forecast the electric load of production machines based on regression algorithms is presented in this paper. The approach is tested and validated on two machine tools, a cleaning machine and an annealing oven of the ETA research factory at Technische Universität Darmstadt [7].

2. LOAD FORECASTING

Load forecasting is a systematic procedure for making statements about future energy demands [8]. In general, load forecasting can be realized with model-driven and data-driven approaches. Model-driven approaches comprise physical and statistical models. Regression analysis [9] and time series analysis [10] represent some of various statistical approaches to load forecasting. On the contrary, machine learning is a data-driven approach to load forecasting [11]. In the literature, the assets and drawbacks of these different approaches have been comprehensively studied, reviewed and compared. Regarding load forecasting, there is no inherently superior method. The approach rather has to be determined based on the specific problem and the availability of data [12].

As of today, electric load forecasting is mainly conducted on the supply side of the energy sector. There, load forecasts are utilized to improve the information base and support the decision making process in the fields of energy purchasing, operations and maintenance or financial planning [13,14]. Although these applications already demonstrate the benefits of load forecasting, the potentials are underutilized in the manufacturing industry. So far, the few existing approaches to forecast a production machine's electric load usually analyze the energy demand of the machines states – off, standby, ready for processing and processing – on the one hand and the different processing steps such as handling, tool exchange or welding on the other hand. Each state and processing step gets an average load assigned which is then recombined for the forecast depending on the production process [15,16]. However, these approaches represent highly simplified forecasts based on average power consumptions not taking further influencing factors into account.

3. SYSTEM AND RESEARCH DEFINITION

Modern production machines offer different machine and process data, such as operating states, states and energy consumption of the machine drives, process specific information like temperatures and pressures, and information about the process steps like the G-code, which contains the manufacturing commands for numerical controlled machine tools. These data sources are utilized on a second based resolution to develop the short term load forecasting approach with a forecasting horizon of 100 seconds. To enable forecasting, the data set has to be prepared depending on the selected forecasting horizon. This data preparation is realized following the procedure from [17]. The approach is developed based on the data of 685 production cycles of a machine tool, which were conducted in seven days. After completion of development the approach is transferred to the data of the remaining production machines based on production data of two days. As the cycle time from the remaining production machines ranges from 110 seconds to 4 hours, the data base for the different systems varies between 4 to 182 cycles.

4. DEVELOPMENT OF MACHINE LEARNING PROCESS FOR LOAD FORECASTING

In order to apply machine learning algorithms, the data has to meet a set of requirements, as some algorithms are not able to deal with missing values. Furthermore, the forecasting accuracy is directly influenced by the data quality and thus by data preprocessing such as missing value imputation, outlier treatment, and data scaling [18]. Therefore, as a first step a preliminary data preprocessing is performed on which a preliminary model selection process is based. Further data preprocessing methods are then evaluated based on the best two learning algorithms. Finally, the best algorithm is selected for further examination of Feature Engineering (FE) strategies in the following step.

As preliminary data preprocessing, instances with more than 50% missing values or missing target values are dropped. Remaining missing values are imputed with the median. For preliminary model selection five promising learning algorithms (Linear Regression, Decision Trees, K-Nearest Neighbor, Random Forest (RF), and Artificial Neural Networks (ANN)) were trained predicting the current electric load. As a result of the preliminary model selection the two best algorithms – RF and ANN – were selected for the examination of further data preprocessing strategies.

First, missing value imputation is analyzed with the mean and median of the respective feature. As outlier treatment analysis a comparison between no treatment at all and outlier treatment with the *Median Absolute Deviation from the Median* (MAD) [19] method was conducted with the threshold set to 4. Furthermore, the numerical features are scaled using the *StandardScaler*, *RobustScaler*, and *MinMaxScaler* from *scikit-learn*, while

Copyright © 2019 ICAE

the categorical features are one-hot-encoded. The evaluation of the performance of those data preprocessing methods is based on the forecasting capability of the two algorithms. As a result of this analysis, the data preprocessing for the forecasting approach has been determined to missing value imputation with the median, no outlier treatment and scaling of the numerical features with the *StandardScaler*. As ANN being the best performing model, it is selected for further investigation.

In a next step, FE in terms of Feature Construction (FC) and Feature Selection (FS) are conducted for the ANN. FS consists of a preliminary selection followed by FC and a final FS. First, all features with a variance of less than 0.1% are eliminated. The small threshold makes sure that a majority of the unnecessary information is eliminated while still making sure that no important information is lost. Furthermore, all features which are highly correlated (Variance Inflation Factor (VIF) >100) are eliminated.

For FC time shift and moving average features are created and combined in three FC strategies which are analyzed in terms of their contribution to the forecasting performance. First, a univariate FC strategy with 100 time shifted and three moving average (10, 30, 60 seconds) features of the electric load. Second, a multivariate FC strategy with 5 time shifted and three moving average (10, 30, 60 seconds) features of each feature and third, a combination strategy with 100 time shifted features of the electric load, one time shifted feature of each feature, and one moving average feature of each feature (60 seconds).

To determine the time shifted features, 1,000 time lags were introduced and the respective number of the best time lags were chosen based on the autocorrelation function. The evaluation of the FC strategies and thus, the decision which strategies to include in the approach is based on the forecasting capability of the resulting feature set. Therefore, the three resulting feature sets are first reduced and subsequently the ANN is hyperparameter tuned utilizing the reduced feature sets. As the combination strategy yielded the worst results, only the two individual strategies are selected.

FS is conducted by recursive feature elimination with a shallow RF with 100 estimators and a maximum depth of 10. The final feature set size is being tuned as a hyperparameter.

For the final hyperparameter tuning of the ANN the network topology is constrained to one hidden layer with different amounts of hidden units. Hyperparameter

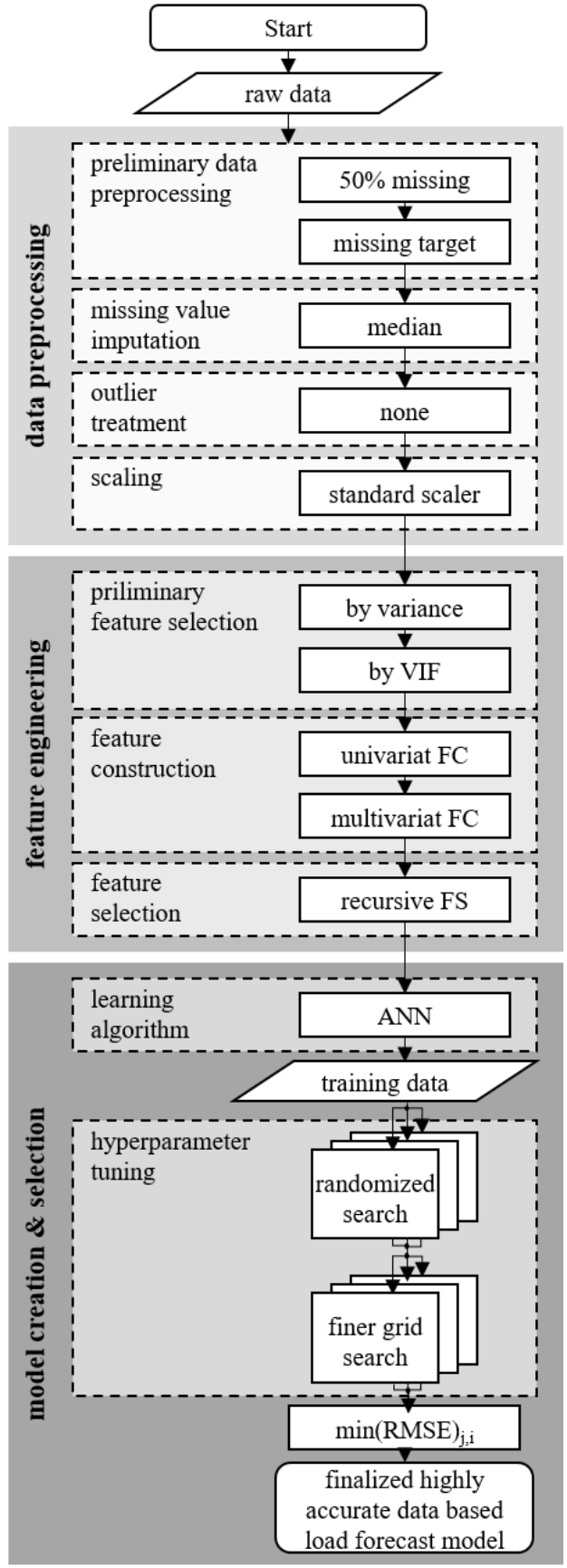

Fig 1: Flow chart of the generic approach

Copyright © 2019 ICAE

Table 1. Grid for Feature Selection and hyperparameter tuning

Step	Parameter	Range
Feature Selection (RF)	number of features	n_features/6 – n_features
	max_features	auto, sqrt
Tuning – random search (ANN)	hidden_layer_sizes	50 – 150
	alpha	3E-06 – 1E-03
	early_stopping	True, false
Tuning – finer gridsearch (ANN)	alpha	(0,37 – 2,72)log(alpha*)
	number of features	(0,6 – 1,4)n_features*

* winning value of random search

tuning is conducted in two steps. First, a randomized coarse search grid is utilized to determine the best hyperparameter combination of the hyperparameters listed in table 1. Subsequently, a finer gridsearch around the yielded best parameter combination is conducted to determine the final model.

A summary of the steps and utilized methods of the forecasting approach is listed in figure 1.

5. APPLICATION AND VALIDATION

Resulting from two different programs per machine tool a total of six models are trained. To validate the results, the scale-independent R^2-score is utilized to compare the results of the different machines. In addition to the training and testing, a baseline score based on a naïve forecast which simply forecasts the current value into the future has been determined. The minimum requirement of the model is to perform better than the naïve forecast. As can be seen in Table 2, this requirement is not only fulfilled, but also greatly exceeded by all models but one. In comparison to the other production machines, the machine tool 2 with program 2 has a complex and highly dynamic load profile, which cannot be captured with the small dataset. Furthermore, the small dataset leads to overfitting, which is an issue for all models of the validation process. Thus, the forecasting capability is expected to improve with increasing amount of training data.

In figure 2 the forecast of the electric load for machine tool 1 on the training and test data set is shown

Table 2. Results of the trained models

Machine	Baseline R^2	Train R^2	Test R^2
Machine Tool 1 – program 1*	0.151	0.861	0.629
Machine Tool 1 – program 2*	0.633	0.744	0.675
Machine Tool 2 – program 1*	0.078	0.800	0.068
Machine Tool 2 – program 2*	0.347	0.830	0.637
Cleaning Machine*	-0.09	0.786	0.547
Annealing Oven°	0.254	0.693	0.481

*Univariate Feature Engineering °Multivariate Feature Engineering

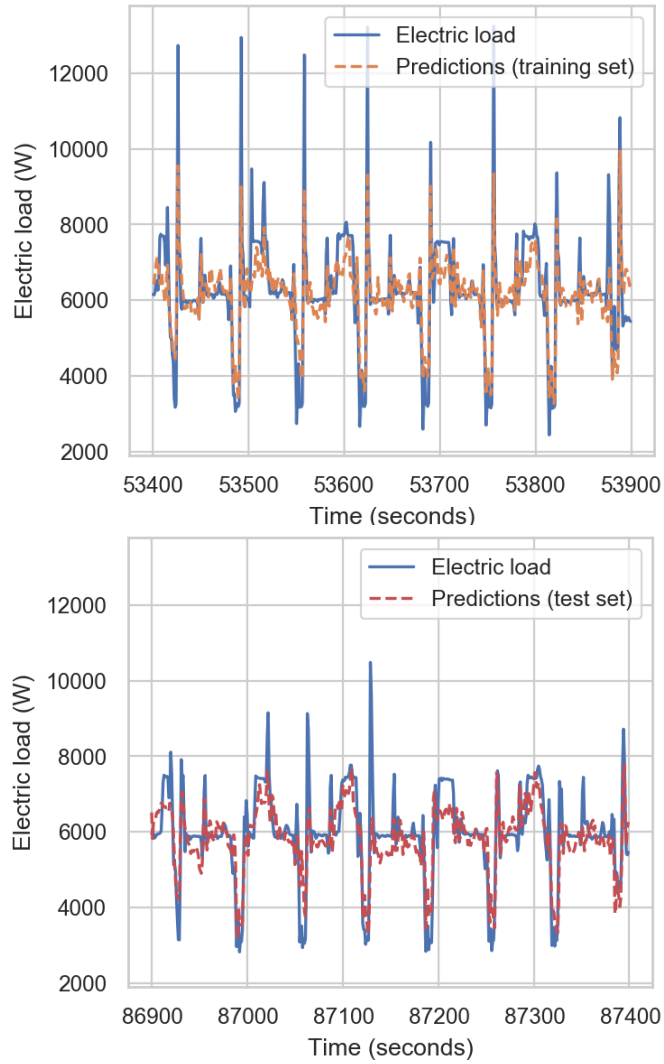

Fig 2: Forcasting of the electric load of machine tool 1

for five production cycles. As can be seen, the forecast captures the general behavior and most peaks well, but seldom the peak height. This applies to the remaining models as well.

6. CONCLUSION

This paper presents a generic machine learning approach to forecast the very short term electric loads of production machines with varying machine and process data. Novelties of the presented approach compared to previous studies are the high manual effort required by those approaches while the presented approach automatically forecasts the electrical load with the usage of Machine Learning. Furthermore, Machine Learning enables the consideration of any number of influencing factors in the model creation, whereas the existing approaches only consider the power consumption, process steps, and operating modes. Here, the seasonal influence of the electrical power consumption of the

Copyright © 2019 ICAE

machine cooling can be mentioned as an example. Moreover, the volatility of the power consumption within individual process steps is captured with the presented approach. In particular, the acceleration and deceleration of the spindle in machine tools lead to high electrical power peaks. As companies not just only pay for the consumed energy, but also for the generated load peaks, only considering average load values per process step, as it is done in the existing studies, leads to the neglect of possibly high energy costs.

Additionally, the presented data based load forecasting approach offers some advantages. First, the low effort and the rapidity to create a forecast. Second, compared to physical modelling a not quite as profound understanding of the system is required. As physical models are often not lean and require a large number of parameters that are difficult to obtain [16].

The results demonstrate that with sufficient data available, the developed machine learning approach is accurate and creates a very short term load forecasting, which can be utilized for various optimization measures to decrease energy costs. First of all, the knowledge of the future energy demand enables optimal energy purchasing. Furthermore, energy costs can be reduced by peak load management through the utilization of electricity storages like battery or flywheel storage systems. Moreover, the information of the energy demand per product can be utilized to implement an energy-adaptive production planning. In addition, the knowledge of the future energy demand enables energy cost savings by demand response applications.

ACKNOWLEDGEMENT

The authors are grateful to the German Federal Ministry for Economic Affairs and Energy (BMWi) for funding the presented work in the project PHI-Factory.

REFERENCES

[1] International Energy Agency. Key World Energy Statistics 2018; 2018.

[2] Statistisches Bundesamt. Umweltnutzung und Wirtschaft-Teil 2: Energie 2000 - 2016. Wiesbaden: Statistisches Bundesamt; 2018.

[3] Hesselbach J, Herrmann C, Detzer R, Martin L, Thiede S, Ludemann B et al. Energy efficiency through optimised coordination of production and technical building services. In: LCE 2008: 15th CIRP International Conference on Life Cycle Engineering: Conference Proceedings; 2008, p. 624.

[4] Alejandro L, Blair C, Bloodgood L, Khan M, Lawless M, Meehan D et al. Global Market for Smart Electricity Meters: Government Policies Driving Strong Growth 2014.

[5] Weinert N, Chiotellis S, Seliger G. Methodology for planning and operating energy-efficient production systems. CIRP annals 2011;60(1):41–4.

[6] Vikhorev K, Greenough R, Brown N. An advanced energy management framework to promote energy awareness. Journal of Cleaner Production 2013;43:103–12.

[7] Panten N, Strobel N, Sossenheimer J, Abele E. Framework for an Energy Efficient and Flexible Automation Strategy and Control Optimization Approach of Supply Systems within a Thermally-Linked Factory. Procedia CIRP 2018;72:526–32.

[8] Clements MP, Hendry DF. Forecasting economic time series. Cambridge: Cambridge Univ. Press; 2000.

[9] Sun X, Ouyang Z, Yue D. Short-term load forecasting based on multivariate linear regression. In: Energy Internet and Energy System Integration (EI2), 2017 IEEE Conference on; 2017, p. 1–5.

[10] Nguyen H, Hansen CK. Short-term electricity load forecasting with Time Series Analysis. In: Prognostics and Health Management (ICPHM), 2017 IEEE International Conference on; 2017, p. 214–221.

[11] Lang K, Zhang M, Yuan Y, Yue X. Short-term load forecasting based on multivariate time series prediction and weighted neural network with random weights and kernels. Cluster Computing 2018:1–9.

[12] Hong T, Fan S. Probabilistic electric load forecasting: A tutorial review. International Journal of Forecasting 2016;32(3):914–38.

[13] Hong T. Short Term Electric Load Forecasting [Dissertation]. Raleigh: North Carolina State University; 2010.

[14] Su P, Tian X, Wang Y, Deng S, Zhao J, An Q et al. Recent Trends in Load Forecasting Technology for the Operation Optimization of Distributed Energy System. Energies 2017;10(9):1303.

[15] Dietmair A, Verl A. Energy consumption forecasting and optimisation for tool machines. Energy 2009;62:63.

[16] Teiwes H, Blume S, Herrmann C, Rössinger M, Thiede S. Energy load profile analysis on machine level. Procedia CIRP 2018;69:271–6.

[17] Walther J, Spanier D, Panten N, Abele E. Very short-term load forecasting on factory level – A machine learning approach. Procedia CIRP 2019;80:705 – 10.

[18] Kotsiantis SB, Kanellopoulos D, Pintelas PE. Data preprocessing for supervised leaning. International Journal of Computer Science 2006;1(2):111–7.

[19] Leys C, Ley C, Klein O, Bernard P, Licata L. Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. Journal of Experimental Social Psychology 2013;49(4):764–6.

Copyright © 2019 ICAE