An Automated Machine Learning Approach for Predicting Chemical Laboratory Material Consumption

António João Silva[1][0000−0003−4747−5073] and Paulo Cortez[1][0000−0002−7991−2090]

ALGORITMI Center, Department of Information Systems, University of Minho,
4804-533 Guimarães, Portugal
id7322@alunos.uminho.pt,pcortez@dsi.uminho.pt

Abstract. This paper address a relevant business analytics need of a chemical company, which is adopting an Industry 4.0 transformation. In this company, quality tests are executed at the Analytical Laboratories (AL), which receive production samples and execute several instrumental analyses. In order to improve the AL stock warehouse management, a Machine Learning (ML) project was developed, aiming to estimate the AL materials consumption based on week plans of sample analyses. Following the CRoss-Industry Standard Process for Data Mining (CRISP-DM) methodology, several iterations were executed, in which three input variable selection strategies and two sets of AL materials (top 10 and all consumed materials) were tested. To reduce the modeling effort, an Automated Machine Learning (AutoML) was adopted, allowing to automatically set the best ML model among six distinct regression algorithms. Using real data from the chemical company and a realistic rolling window evaluation, several ML train and test iterations were executed. The AutoML results were compared with two time series forecasting methods, the ARIMA methodology and a deep learning Long Short-Term Memory (LSTM) model. Overall, competitive results were achieved by the best AutoML models, particularly for the top 10 set of materials.

Keywords: Industry 4.0 · Automated Machine Learning · Regression · Time Series Forecasting · Deep Learning.

1 Introduction

With the emergence of the Industry 4.0 concept, there is an increase of digital transformation, where industrial physical processes generate data that can be analyzed by Machine Learning (ML) algorithms to provide valuable Business Analytics (BA) [19]. These analyses can impact several production aspects, including stock management. However, in the Chemical industry the usage of ML and BA is still scarce.

In this work, we address a BA need of a chemical organization that is adopting an Industry 4.0 transformation in their Analytical Laboratories (AL). During

the production process, selected samples are sent to be tested at the AL, which is responsible for assuring that the products are compliant with quality standards. The analysis of a sample at the AL requires diverse instrumental analyses, each consuming one or more materials (e.g., Acetone, Dichloromethane, Ethanol, Methanol). Under this context, predicting the amount of materials needed for the quality tests is crucial to support a AL stock management, preventing quality inspection delays which would prejudice production. In previous work [20], we have adopted a ML approach to successfully predict the arrival times of samples at the AL. By using this predictive approach, the chemical organization can now perform weekly plans of the expected instrumental AL usage. Under this context, this paper describes a ML approach to predict the weekly consumption of AL materials based on the expected instrument usage. The approach was developed using the CRoss-Industry Standard Process for Data Mining (CRISP-DM) methodology [26]. Similarly to the work conducted in [20], to better focus on feature engineering (data preparation phase of CRISP-DM), we adopt an Automated Machine Learning (AutoML) [8], which is executed during the modeling CRISP-DM phase and that allows to automatically select and tune the hyperparameters of the predictive ML models. Using real-word data, collected from a chemical company, we executed several CRISP-DM iterations, exploring three main input variable selection strategies and two sets of AL materials (top 10 and all consumed materials). The experimentation adopts a realistic rolling window evaluation scheme, which simulates several train and test modeling updates through time. For benchmark purposes, the proposed ML approach is compared with two time series forecasting methods: the known ARIMA methodology [1] and a deep learning Long Short-Term Memory (LSTM) [17].

The paper is structured as follows. Section 2 describes the related work. The problem contextualization is presented in Section 3. Next, the analyzed data and prediction methods are presented in Section 4. Then, the obtained results are shown and discussed in Section 5. Finally, Section 6 concludes the paper.

2   Related Work

The Industry 4.0 concept [19] is impacting diverse industrial sectors. With the increased usage of interconnected sensors (e.g., Internet-of-Things), factories generate more digital data that reflect their production processes. All these data can be analyzed by AI and ML algorithms, providing valuable BA. In some cases, real-world ML project fail due to to a misalignment between business needs and ML analyses [7]. The CRISP-DM methodology was precisely developed to solve this issue, increasing the success of ML projects [26]. The methodology involves both business and ML experts and includes six main phases: business understanding, data understanding, data preparation, modeling, evaluation and deployment. In previous works, we have employed CRISP-DM to successfully model the business needs of textile [18] and chemical [20] companies.

Turning to the specific chemical sector, in several organizations the Industry 4.0 concept is not yet fully embraced. For instance, while quality tests are rigor-

ously stored in digital databases, the same does not occur with the AL processes [21, 13]. Also, it is common to have information silos (e.g., production, AL) and thus this lack of database interoperability diminishes the full potential to use ML to extract valuable BA from the data [20]. Thus, most predictive analytics studies for the chemical sector involve the production processes, rather than AL (such as executed in this paper). For instance, Roe et al. [25] used a Fuzzy Neural Network model to perform a predictive control on a solar-thermal chemical processing. Moreover, Longone et al. [14] used a Logistic Regression to predict production anomalies in a chemical plant that adopted the Industry 4.0 concept. It should be highlighted that most predictive ML studies in industry are focused on non chemical sectors and target the predictive maintenance task. Examples of ML algorithms that were proposed for such task include: Random Forest (RF) [3], Neural Networks (NN) [22], Gradient Boosting Machines (GBM) [15] and Support Vector Machines (SVM) [23]. In all these ML predictive studies, expert knowledge and trial-error experiments were used to select and tune the predictive ML algorithms, which is a common ML practice. However, there is a recent ML trend that assumes the usage of AutoML [8]. The main advantage of AutoML is that it alleviates the ML analyst effort, allowing to focus on other aspects of the ML pipeline process (e.g., data engineering). In [20], we have adopted an AutoML approach to predict the arrival of production samples at the AL, allowing to support the allocation of human resources and analytical equipment. In this paper, we adopt a similar AutoML approach but focusing on a different business need from the same chemical company: to predict the week AL material consumption based on quality instrumental usage estimates.

3   Problem Formulation

Figure 1 presents the flow of main transactions that occur between three main sections of the analyzed chemical company: Warehouse, Production and Analytical Laboratories (AL). The Warehouse is responsible for storing and managing the different materials that are provided by the suppliers and that are needed by the company. (e.g., raw production materials). In this work, we focus on analytical materials, which are used in the AL. The Production line is where the production process is performed. A production of a certain product starts when there is a production order for that product on that specific date. A production order contains the several informational elements: the product to be produced, the quantity in batches to be produced, the raw materials to be used and the start and end dates. The dates are added to the database when the production ordered ends. During the production period, several production samples, called In Production (IP), are sent to the AL for quality assessment. If quality is below the client requirements, then the production line will have to perform adjustments, in order improve the expected quality of the product. Thus, the AL are a critical element of the production process, with delays in AL testing resulting in production stops and delays in the execution of new production orders.
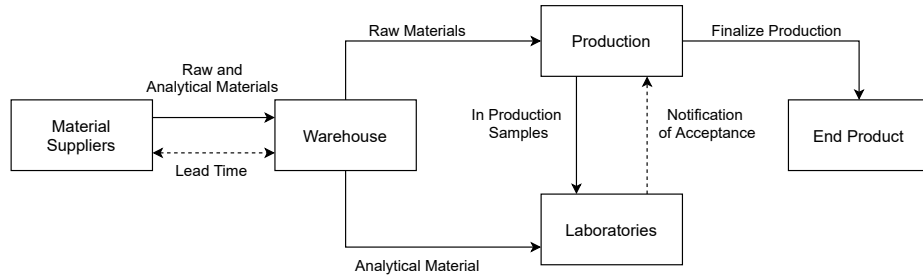
Fig. 1. Workflow of materials and production transactions.

At the AL, the quality tests use several instrumental analyses that require analytical materials, in order to guarantee the feasibility of the tests. When there is an AL shortage of materials, they are ordered from the Warehouse, using the Enterprise Resource Planning (ERP) production system. In some cases, there is a low stock of the analytical materials in the Warehouse, which needs to produce supplier orders that take time, thus producing AL quality testing delays. In previous work [20], we have adopted an AutoML approach to predict the arrival of IP production samples at the AL. Using such predictions, the company information system is capable of producing accurate week plans of AL instrumental needs. In this paper, the ML goal is to use the AL tests (or plans) as the inputs of a regression model, aiming to predict a particular analytical material consumption. Let $\mathbf{X}$ denote a data matrix $N \times Q$ with the elements $x_{i,j}$, each representing the number of quality tests of type j that were executed (or are planned) for a particular week i, where N is the total number of weeks and Q is the total number of distinct quality tests. Let $\mathbf{Y}$ denote a matrix $N \times M$ with the elements $y_{i,m}$, each representing the quantity of consumed material of type $m \in \mathcal{M}$ for the week i, where $\mathcal{M} = \{1, 2, ..., M\}$ denotes a selection set with M distinct analytical materials. Another relevant business concept is the AL total weekly consumption quantity ($T_{\mathcal{M}}$), computed as $T_{\mathcal{M}} = \sum_{m=1}^{M} y_{i,m}$. The total consumption quantity is useful for resizing the AL warehouse.

The business goal is to estimate the w weekly quantity $\hat{y}_{w,m}$ based on the quality tests that use the m material:

$$\hat{y}_{w,m} = f(x_{w,k_1}, ..., x_{w,k_K}) \tag{1}$$

where $\{k_1, ..., k_K\}$ denotes the set of laboratory tests that are used as inputs and f is the data-driven function that will be learned using the AutoML approach. In this work, each m material consumption prediction requires the training of a different ML model. Moreover, the $\{k_1, ..., k_K\}$ input tests are dependent of the adopted feature selection strategy (Section 4.2). Once the distinct ML predictive models are built, the AL total weekly consumption quantity for selection $\mathcal{M}$ can be computed as: $\hat{T}_{\mathcal{M}} = \sum_{m=1}^{M} \hat{y}_{w,m}$.

4 Materials and Methods

4.1 Data

The data used in this study was retrieved by executing an Extract, Transform & Load (ETL) process, which extracted data records from two main databases related with the production and AL units. The resulting dataset includes a total of $N = 177$ weeks of data, from January 2016 to May 2019. In total, the input X matrix includes a total of $Q = 30$ distinct quality tests, thus with 177×30 elements. Some of the analyzed input tests have a strong correlation, while other variables often include a large portion of zero values. In Section 4.2, we will use these properties to design feature selection strategies. As for the target Y matrix, it includes a total of $M = 26$ analytical materials (e.g., Acetone, Ethanol, Methanol) After consulting the company experts, we explore two main sets of prediction targets: top 10 - with the $M = 10$ highest consumed materials ($\mathcal{M} = \{1, ..., 10\}$); and all – with all $M = 26$ materials ($\mathcal{M} = \{1, ..., 26\}$). Due to commercial privacy concerns, we do not disclose further details about the specific analyzed variables.

4.2 Prediction Methods

We adopted the `R` computational tool and its `rminer` package [6] for data manipulation and computation of the ML regression metrics. The AutoML is based on the `H2O` open-source tool (https://www.h2o.ai/products/h2o-automl/) [5]. The `auto.arima` from the `forecast` `R` package was used to automate and fit the ARIMA models [1, 11, 12]. Finally, the LSTM model was implemented using the `PyTorch` Python module [17].

The AutoML models were configured to select the best regression model and its hyperparameters for each targeted m material. The selection is based on the best Root Mean Squared Error (RMSE) computed using a validation set that is obtained by applying an internal 10-fold cross-validation method over the training data. All computational experiments were executed on the same personal computer and each individual ML model was trained up to a maximum running time of 3,600 seconds. Once a ML model is selected, the model was retrained with all training data. As in [8], the AutoML was configured to include a total of 6 distinct regression algorithms: RF, Extremely Randomized Trees (XRT), Generalized Linear Model (GLM), GBM, XGBoost (XG) and a Stacked Ensemble (SE). The RF is a popular ensemble method that combines a large number of decision trees based on bagging and random selection of input features [10]. The XRT algorithm extends the RF approach by randomly selecting the decision thresholds of the tree nodes [9]. GLM estimates regression models for exponential distributions (e.g., Gaussian, Poisson, gamma) [10]. The GBM algorithm is a based on a generalization of tree boosting, sequentially building regression trees for all data features [10]. XG is another ensemble tree method that uses boosting to enhance the prediction results [4]. The SE method, also known as

stacked regression [2], combines the predictions of different base learners by using a second-level ML algorithm. The H2O implementation [5] uses the following AutoML setup: RF and XRT – set with the default hyperparameters; GLM - grid search used to set one hyperparameter (*alpha*, a regularization parameter); GBM and XG – grid search used to tune nine and ten hyperparameters (e.g., number of trees, maximum depth, minimum rows); SE – all five algorithms (RF, XRT, GLM, GBM, XG) are used as base learners and the individual predictions are weighted by using a second-level GLM learner.

The input matrix $\mathbf{X}$ includes several variables that are either correlated with other variables or contain a large number of zero values. In order to improve the AutoML results, we explore three main input Feature Selection (FS) strategies, that were applied to the training data: ALL - with all $Q = 30$ inputs, executed during the first CRISP-DM iteration; FS1 – all variables with a correlation higher than 60% or with more than 90% of zeros are removed (resulting in $Q = 15$), executed during the second CRISP-DM iteration; and FS2 - all variables with a correlation higher than 90% or with more than 90% of zeros are removed (leading to $Q = 19$), executed during the first CRISP-DM iteration.

For comparison purposes, we also consider two main time series forecasting methods, each using only the $y_{i,m}$ past observations ($i \in \{1, ..., m-1\}$) to predict $\hat{y}_{w,m}$ at week w: ARIMA and LSTM. The ARIMA is automatically build using the `forecast` R package, while the LSTM assumes a default parametrization with one input node (first time lag, y_{t-1}, where t is the current time), one hidden layer with 100 hidden nodes and hyperbolic tangent activation function, one output node (current observation, y_t), the Adam optimizer, Mean Squared Error (MSE) loss function and 150 training epochs.

4.3   Evaluation

We adopted a Rolling Window (RW) evaluation scheme [24, 16], which simulates a realistic execution of the AutoML models by performing several training and test updates through time (Figure 2). With this scheme, the initial training set with a fixed size of W time periods is used to generate the training models and execute a one week ahead prediction ($T = 1$). Then, the W data is updated by discarding the oldest week observations and adding one subsequent week of data. A new prediction model is built, allowing to issue a new prediction, and so on. In total, the RW results in $U = N - W$ training and testing updates. In this work, we have set $W = 147$, which allows to obtain $U = 30$ RW iterations. In order to reduce the computational effort, since we conduct a large number of ML experiments (e.g., we target $M = 26$ distinct outputs), the AutoML model and hyperparameter selection is only executed once for each m material, using the training data from the first RW iteration. Once the ML is selected, it is retrained for each RW iteration.

As for the regression metrics, using the $U = 30$ test predictions, we compute five measures [10, 16]: Mean Absolute Error (MAE), Normalized MAE (NMAE), RMSE, Relative Squared Error (RSE), and the coefficient of determination (R^2). The lower the MAE, NMAE and RMSE values the better are the predictions. The
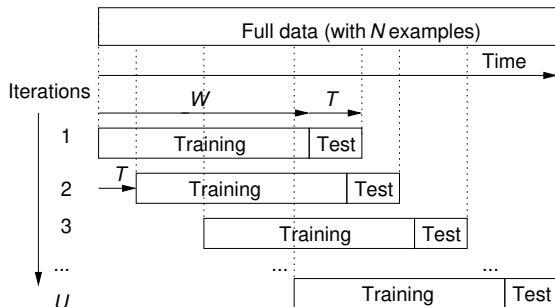
Fig. 2. Schematic of the Rolling Window (RW) evaluation.

NMAE measure is computed as $\frac{MAE}{\max(y_{i,m})-\min(y_{i,m})}$, where $y_{i,m}$ denotes the target variable for material m. When compared with MAE, the NMAE metric presents two main advantages [16]: it is more easy to interpret, since it expresses the error as a percentage of the full target scale (y); it is scale independent, which is useful for the analytical consumption data given that we handle different materials and thus distinct consumption scales. The RMSE measure is particularly important in this domain, since it is more sensitive to extreme values when compared with MAE. Thus, a lower RMSE should be aligned with a better upper or lower peak prediction, which is more useful to assist the stock management of the consumed AL materials. The RSE is computed as $\frac{SSE_{\hat{y}_{i,m}}}{SSE_{\overline{y}_{i,m}}}$, where SSE denotes the sum of squared errors and $\overline{y}_{i,m}$ the average of the target variable on the test data. The RSE is similar to the RMSE measure in the sense that it is also more sensitive to extreme errors. The advantage is that RSE is scale independent. While the RSE values can be also presented as percentages (such as NMAE), the RSE values are more difficult to interpret by end users, since it only expresses how good are the predictions when compared with the average target values. As for R^2, it measures the goodness of fit. The higher value, the better is the alignment between consecutive changes in the predicted and real values, with the perfect regression model producing a maximum of $R^2=1$.

Since we target a large number of individual models (up to $M = 26$), the value of each forecasting approach is globally measured by considering the predictive measures applied to total quantity consumption target for a particular $\mathcal{M}$ selection. For instance, the RW MAE is computed as $MAE = \sum_{u_1}^{U} |T_{\mathcal{M}} - \hat{T}_{\mathcal{M}}|/U$, where u is a RW iteration and $\hat{T}_{\mathcal{M}}$ is the predicted total quantity consumption.

5   Results and Discussion

Table 1 summarizes the obtained RW predictive results for the total quantity consumption and $\mathcal{M}$ selection of materials. For instance, the upper left value of 193.0 corresponds to the MAE average when considering all $m \in \mathcal{M}, \mathcal{M} = \{1, 2, ..., 10\}$ highest consumed analytical materials of the top 10 selection set.

The results from Table 1 confirm that different CRISP-DM iterations produced improved predictions, with the FS2 feature selection strategy obtaining the best AutoML results for all regression metrics. As for the time series forecasting baselines, the ARIMA methodology outperformed the LSTM neural network approach. Overall, the AutoML FS2 method produces the best predictions for the top 10 selection (for all regression measures) and the best RMSE, RSE and R^2 values for the all selection ($M = 26$). As explained in Section 4.3, for the improving stock management of the analytical materials, the squared error measures (RMSE and RSE) are more important than absolute error ones (MAE and NMAE). Regarding the optimized ML models, the AutoML procedure selected only three of the six considered regression algorithms: GLM, GBM and RF.

Table 1. Summary of the RW predictive results (best values in **bold**).

Method	Top 10 ($M = 10$)					All ($M = 26$)				
	MAE	NMAE	RMSE	RSE	R^2	MAE	NMAE	RMSE	RSE	R^2
AutoML ALL	193.0	6.30	338.0	51.9	0.49	80.58	2.63	205.6	42.4	0.58
AutoML FS1	203.7	6.66	347.4	54.9	0.46	83.86	2.74	209.6	44.1	0.56
AutoML FS2	**187.7**	**6.13**	**330.2**	**49.5**	**0.51**	78.89	2.58	**200.8**	**40.5**	**0.60**
ARIMA	189.1	6.18	349.0	55.3	0.47	**76.92**	**2.51**	210.4	44.5	0.57
LSTM	230.0	7.52	367.7	61.4	0.41	90.67	2.96	219.1	48.2	0.53

For demonstration purposes, Figure 3 shows the RW predictions for the selected AutoML FS2 method, which provided the lowest squared errors and highest coefficient of determination values. Due to business privacy issues, the scale values of the y-axis are omitted from the plots. In the plots, we also present in brackets the NMAE errors, since these are more easy to be interpreted by the chemical experts. The top two graphs show the results when predicting the total consumption (top 10 or all), while the middle and bottom graphs denote the prediction results for four individual materials ($m \in \{2, 10, 13, 17\}$). Overall, the real and predictive curves are very close and the prediction models are capable of correctly identifying several high and low consumption peaks, thus confirming that high quality predictions were obtained by the AutoML FS2 method.

The obtained results were shown to the chemical company experts, which highlighted the total quantity results, which can be used to resize the AL warehouse. Moreover, the chemical experts considered that individual material predictions are interesting, such as for $m = 2$ and $m = 17$ from Figure 3, which have a strong potential to improve the stock management of these materials.

6   Conclusions

This study addresses a relevant business goal of a chemical company that is being transformed under the Industry 4.0. In particular, a Machine Learning
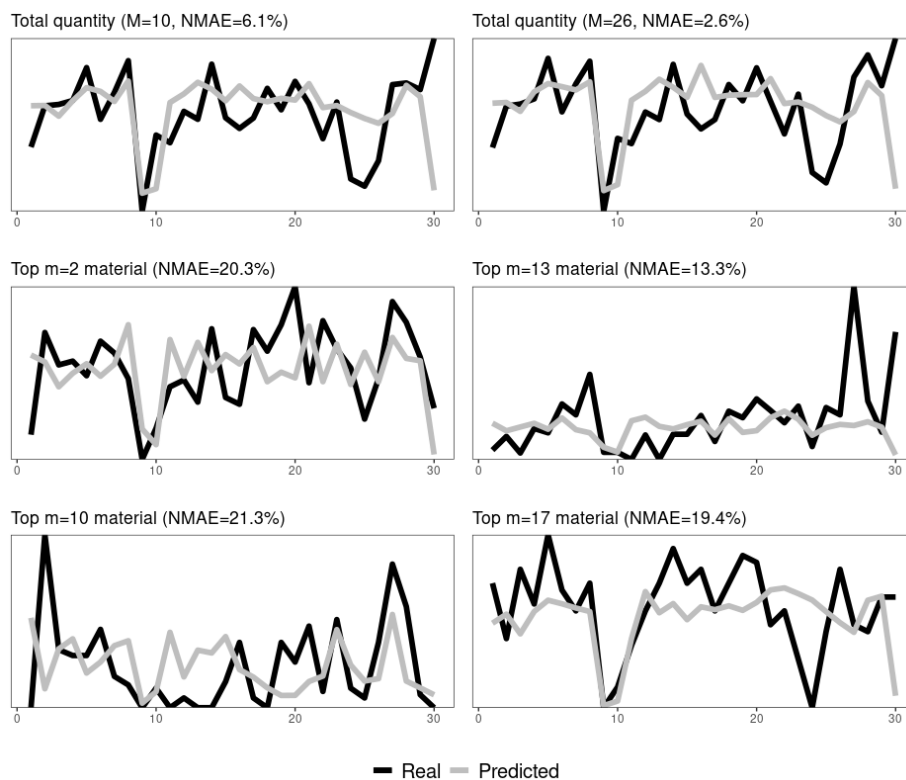
Total quantity (M=10, NMAE=6.1%)

Total quantity (M=26, NMAE=2.6%)

Top m=2 material (NMAE=20.3%)

Top m=13 material (NMAE=13.3%)

Top m=10 material (NMAE=21.3%)

Top m=17 material (NMAE=19.4%)

■ Real ■ Predicted

Fig. 3. RW predictive results for AutoML FS2 method (x-axis denotes the considered week, from March 2019 to May 2019; y-axis shows the analytical material consumption).

(ML) approach was conducted, aiming to predict the needs of materials (e.g., Acetone, Ethanol) used in their Analytical Laboratories (AL). The ML project was conducted using the CRoss-Industry Standard Process for Data Mining (CRISP-DM) methodology. At the data understanding CRISP-DM stage, we collected 177 weeks of data, from January 2016 to May 2019, involving a total of 30 quality tests and up to 26 consumed AL materials. It should be noted that the chemical company is currently capable of producing weekly quality test usage plans with a good accuracy. Thus, the regression goal is to model AL material consumption as a function of the conducted quality tests. Using the collected data, we have developed large set of regression models (total of $M =26$ models), which were analyzed in terms of two major sets of material selections: top 10 most consumed materials ($M=10$) and all materials ($M=26$). To reduce the ML analyst effort, we have employed an Automated Machine Learning (AutoML) procedure during the CRISP-DM modeling stage, which allows to automatically select the best among six different regression algorithms. A total of three CRISP-DM iterations were executed, each exploring a different Feature Selection (FS) method. For comparison purposes, we also considered two time series forecasting methods: ARIMA and a Long Short-Term Memory (LSTM) neural network.

Several computational experiments were executed, by considering a realistic Rolling Window (RW) procedure that simulated 30 training and testing iterations through time. The best overall results were achieved by the AutoML FS2 method (corresponding to the third CRISP-DM iteration), which obtained a total quantity Normalized Mean Absolute Error (NMAE) of 6.1% (top 10 selection) and 2.6% (all materials). The predictive results were shown to the AL managers, which provided a positive feedback. Indeed, in future work, we intend to focus on the development stage of CRISP-DM, deploying the studied prediction models in the chemical company information system. This will allow to measure the business value of using these predictions to improve the warehouse stock management of analytical materials.

Acknowledgments

References

1. Box, G.E.P., Pierce, D.A.: Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. Journal of the American Statistical Association **65**(332), 1509–1526 (1970), http://www.jstor.org/stable/2284333
2. Breiman, L.: Stacked regressions. Machine Learning **24**(1), 49–64 (1996). https://doi.org/10.1007/BF00117832

3. Canizo, M., Onieva, E., Conde, A., Charramendieta, S., Trujillo, S.: Real-time predictive maintenance for wind turbines using big data frameworks. In: 2017 IEEE International Conference on Prognostics and Health Management, ICPHM 2017, Dallas, TX, USA, June 19-21, 2017. pp. 70–77. IEEE (2017). https://doi.org/10.1109/ICPHM.2017.7998308

4. Chen, T., Guestrin, C.: Xgboost: A scalable tree boosting system. In: Krishnapuram, B., Shah, M., Smola, A.J., Aggarwal, C.C., Shen, D., Rastogi, R. (eds.) Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, August 13-17, 2016. pp. 785–794. ACM (2016). https://doi.org/10.1145/2939672.2939785

5. Cook, D.: Practical machine learning with H2O: powerful, scalable techniques for deep learning and AI. " O'Reilly Media, Inc." (2016)

6. Cortez, P.: Modern optimization with R. Springer (2014)

7. Deal, J.: The ten most common data mining business mistakes (Jun 2013), https://www.elderresearch.com/most-common-data-science-business-mistakes

8. Ferreira, L., Pilastri, A., Martins, C., Santos, P., Cortez, P.: An automated and distributed machine learning framework for telecommunications risk management. In: Rocha, A.P., Steels, L., van den Herik, H.J. (eds.) Proceedings of the 12th International Conference on Agents and Artificial Intelligence, ICAART 2020, Volume 2, Valletta, Malta, February 22-24, 2020. pp. 99–107. SCITEPRESS (2020). https://doi.org/10.5220/0008952800990107, https://doi.org/10.5220/0008952800990107

9. Geurts, P., Ernst, D., Wehenkel, L.: Extremely randomized trees. Machine Learning **63**(1), 3–42 (2006). https://doi.org/10.1007/s10994-006-6226-1

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

11. Hyndman, R., Athanasopoulos, G., Bergmeir, C., Caceres, G., Chhay, L., O'Hara-Wild, M., Petropoulos, F., Razbash, S., Wang, E., Yasmeen, F.: forecast: Forecasting functions for time series and linear models (2020), https://pkg.robjhyndman.com/forecast/, r package version 8.13

12. Hyndman, R.J., Khandakar, Y.: Automatic time series forecasting: the forecast package for R. Journal of Statistical Software **26**(3), 1–22 (2008), https://www.jstatsoft.org/article/view/v027i03

13. Kammergruber, R., Robold, S., Karliç, J., Durner, J.: The future of the laboratory information system–what are the requirements for a powerful system for a laboratory data management? Clinical Chemistry and Laboratory Medicine (CCLM) **52**(11), 225–230 (2014)

14. Langone, R., Cuzzocrea, A., Skantzos, N.: Interpretable anomaly prediction: Predicting anomalous behavior in industry 4.0 settings via regularized logistic regression tools. Data & Knowledge Engineering **130**, 101850 (2020). https://doi.org/https://doi.org/10.1016/j.datak.2020.101850

15. Liulys, K.: Machine learning application in predictive maintenance. In: 2019 Open Conference of Electrical, Electronic and Information Sciences (eStream). pp. 1–4. IEEE (2019)

16. Oliveira, N., Cortez, P., Areal, N.: The impact of microblogging data for stock market prediction: Using twitter to predict returns, volatility, trading volume and survey sentiment indices. Expert Syst. Appl. **73**, 125–144 (2017). https://doi.org/10.1016/j.eswa.2016.12.036, https://doi.org/10.1016/j.eswa.2016.12.036

17. Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Kopf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., Bai, J., Chintala, S.: Pytorch: An imperative style, high-performance deep learning library. In: Wallach, H., Larochelle, H., Beygelzimer, A., d'Alché Buc, F., Fox, E., Garnett, R. (eds.) Advances in Neural Information Processing Systems 32, pp. 8024–8035. Curran Associates, Inc. (2019), http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf

18. Ribeiro, R., Pilastri, A., Moura, C., Rodrigues, F., Rocha, R., Morgado, J., Cortez, P.: Predicting physical properties of woven fabrics via automated machine learning and textile design and finishing features. In: Maglogiannis, I., Iliadis, L., Pimenidis, E. (eds.) Artificial Intelligence Applications and Innovations - 16th IFIP WG 12.5 International Conference, AIAI 2020, Neos Marmaras, Greece, June 5-7, 2020, Proceedings, Part II. IFIP Advances in Information and Communication Technology, vol. 584, pp. 244–255. Springer (2020). https://doi.org/10.1007/978-3-030-49186-4_21

19. Shrouf, F., Ordieres, J., Miragliotta, G.: Smart factories in industry 4.0: A review of the concept and of energy management approached in production based on the internet of things paradigm. In: 2014 IEEE International Conference on Industrial Engineering and Engineering Management. pp. 697–701 (Dec 2014). https://doi.org/10.1109/IEEM.2014.7058728

20. Silva, A.J., Cortez, P., Pilastri, A.: Chemical laboratories 4.0: A two-stage machine learning system for predicting the arrival of samples. In: Maglogiannis, I., Iliadis, L., Pimenidis, E. (eds.) Artificial Intelligence Applications and Innovations - 16th IFIP WG 12.5 International Conference, AIAI 2020, Neos Marmaras, Greece, June 5-7, 2020, Proceedings, Part II. IFIP Advances in Information and Communication Technology, vol. 584, pp. 232–243. Springer (2020). https://doi.org/10.1007/978-3-030-49186-4_20, https://doi.org/10.1007/978-3-030-49186-4_20

21. Skobelev, D., Zaytseva, T., Kozlov, A., Perepelitsa, V., Makarova, A.: Laboratory information management systems in the work of the analytic laboratory. Measurement Techniques **53**(10), 1182–1189 (2011)

22. Spendla, L., Kebisek, M., Tanuska, P., Hrcka, L.: Concept of predictive maintenance of production systems in accordance with industry 4.0. In: 2017 IEEE 15th International Symposium on Applied Machine Intelligence and Informatics (SAMI). pp. 000405–000410. IEEE (2017)

23. Straus, P., Schmitz, M., Wostmann, R., Deuse, J.: Enabling of predictive maintenance in the brownfield through low-cost sensors, an iiot-architecture and machine learning. In: 2018 IEEE International Conference on Big Data (Big Data). pp. 1474–1483 (12 2018). https://doi.org/10.1109/BigData.2018.8622076

24. Tashman, L.J.: Out-of-sample tests of forecasting accuracy: an analysis and review. International Journal of Forecasting **16**(4), 437 – 450 (2000). https://doi.org/10.1016/S0169-2070(00)00065-0

25. Wen, Z., Xie, L., Fan, Q., Feng, H.: Long term electric load forecasting based on ts-type recurrent fuzzy neural network model. Electric Power Systems Research **179**, 106106 (2020). https://doi.org/https://doi.org/10.1016/j.epsr.2019.106106

26. Wirth, R., Hipp, J.: Crisp-dm: Towards a standard process model for data mining. In: Proceedings of the 4th international conference on the practical applications of knowledge discovery and data mining. pp. 29–39 (2000)