XAI FOR ANOMALY ANALYSIS BY POWER PLANT OPERATORS - A CASE AND USER STUDY

Marcel Dix[1], **Jan Koltermann**[2], **Sebastian Mieck**[2], **Boris Pastler**[1], **and Benjamin Kloepper**[1,3]
[1]Industrial AI, ABB AG Corporate Research Center, Germany
[2]Power Plant Analytics, LEAG Lausitz Energy Power Plants AG, Germany
[3]Intelligent Industry, Capgemini Invent, Germany

ABSTRACT

Detecting and analyzing abnormal situations in power plants is a crucial task for plant operators. This paper explores the potential of Explainable AI (XAI) techniques in supporting operators by providing comprehensive insights into the origins and causes of anomalies. A case study was conducted at the LEAG power plant in Boxberg, Germany — one of Europe's largest power plants. The study utilized a dashboard developed in this study to present AI-predicted anomalies to operators. To enhance the plausibility of anomalies and understanding of their root causes, the dashboard incorporates five basic explanatory techniques from the XAI literature: case-based explanations, feature importance, counterfactual explanations, contextualization, and natural language explanations. The case study includes a user evaluation with power plant operators from LEAG to assess the effectiveness of XAI techniques in enhancing operators' analysis and understanding of abnormal situations. The results confirm the value of these techniques, and their implications are discussed in this paper.

Keywords AI, Explainable AI (XAI), Anomaly Detection, Power Plant, Operator

1 Introduction

The growing trend of digitalization in power plants results in the continuous increase of vast amounts of data — primarily time series data from hundreds of digital sensors in the plant. The volume of historical production data being processed and stored can easily accumulate to several terabytes. All this data represents a great potential for novel AI use cases that could contribute to more efficient and safer operation of power plants. Here, a relevant AI use case often considered in literature is to assist plant operators in the detection of abnormal situations in the plant [1, 2]. Once an abnormal situation is detected, the operator must investigate the situation, identify its cause, and determine an appropriate course of action in response to the situation. This part of the operator's job is particularly challenging: It requires the observation of dozens or even hundreds of signals spread across multiple plant sections and operator screens, and relies heavily on the operator's personal experience and current mental freshness.

An important limitation of AI is that many machine learning models lack transparency to explain why the model reached a particular prediction [3]. To overcome this limitation, the research field “Explainable AI” (XAI) [4, 5] has received increasing attention also in the industrial domain recently. This paper presents a case and user study of how XAI can support the evaluation, analysis, and understanding of potentially abnormal situations detected by a machine learning-based anomaly detection method. This approach goes beyond the common use of XAI to provide insight into the reasoning process of machine learning models, but investigates whether the XAI output also provides valuable insight into the current process conditions. For this study, a dashboard has been developed at ABB Corporate Research Germany, as part of the EU-funded research project EXPLAIN [6], and in collaboration with LEAG. In the provided user scenario, the dashboard serves as an AI-based operator support system, e.g. as an additional operator screen in the control room, where anomalies, detected by an underlying anomaly detection model, are displayed, and the model output is explained to the operator. The explanation of anomalies is realized with the help of serveral visual *explainers* that have been developed in this study, based on five underlying XAI techniques from literature, namely: showing similar anomalies (case-based explanations), showing the signal contributions to an anomaly (feature importance), showing the actual vs. expected signal trend (counterfactual explanations), showing related alarms (contextualization), and generating text-based explanations for the numerical data (natural language explanations). In the evaluation with several power plant operators, we are able to demonstrate the explanatory value of the presented solution for this industrial domain.

2 Related Work

The debate in literature about "Explainable AI" (XAI) concerns the need that predictions from AI models appear plausible to the user. Some ML models, such as linear models and decision trees, are inherently interpretable [7]. On the other hand, many ML models, and here most notably deep neural networks, present themselves rather as opaque black boxes to the user [4]. Various XAI techniques have been built to provide insights into the general reasoning of ML models or explaining specific model predictions. The general focus of XAI is to increase trust and to enable better supervision and debugging of ML models. In our case study, we extend this scope of XAI and investigate, whether XAI appplied to ML models using process data as inputs can help process operators (specifically power plant operators) to gain insights into the condition of the industrial processes and support them in the task of analyzing abnormal situations.

An important categorization for XAI techniques is by the type of explanations they generate. Here, feature importance methods [8, 9], prototypes (creation of a prototype, illustrating how a class typically looks [10]), case-based explanations (e.g., finding similar cases in the training dataset [11]), contextualization (explaining by adding, e.g., domain-specific context information) [12], natural language explanations (i.e. rephrasing a model output in natural language) [13], causal mechanisms (e.g., rule generation [11]), and counterfactual explanations [14] can be distinguished. The proposed solution in this paper presents an example implementation for five of the above-named XAI techniques, as mentioned in the introduction.

Many of these XAI techniques are limited to specific data types, especially tabular, image, and textual data [15]. While these data types can be found in power plants, too (e.g., in the form of tabular alarm logs, IR thermal images for equipment monitoring, or textual data in electronic operator shift books), the monitoring of plants is primarily based on *time series data* (i.e. sensor readings from the automation equipment in the plant). After initially receiving little attention in XAI research [16, 17], time series data has recently gained increasing attention, with more publications appearing on the topic [18]. For example, in 2017, LIME, originally developed by Ribeiro et al. [8] to explain tabular, textual, and image data, was extended by Metzenthin [19] to include time series data, resulting in the creation of the new library LIME-for-Time. Another popular technology that has been applied to time series is SHAP [9], but both LIME-for-Time and SHAP suffer from being limited to univariate time series data [20]. Indeed, a notable part of prior art we found is limited to explain univariate time series, whereas many industrial AI use cases, including the one discussed in this paper, involve *multivariate* time series. Furthermore, we noted a lack of literature providing industrial *examples* [21]. This study aims to contribute towards filling these gaps.

3 Case Study

The LEAG power plant in Boxberg, Germany, commenced operations in the 1970s, and today consists of four plant units with a total capacity of 2575 megawatts (gross). The specific unit considered in this study is called "Unit Q" which has a capacity of 900 megawatts (gross) [22]. Operations began in 2000 and are monitored around the clock from the unit's central control room by a team of 4-5 operators. The operators work in rotating shifts of 8-12 hours. A typical shift is characterized by daily routines with extended periods of inactivity for the operator. Operators would describe a "good shift" as one with little to no activity. However, the complexity and amount of activity sharply increases during abnormal situations, and then, the operator's experience plays a critical role in identifying and resolving these anomalies. Here, an AI-based operator support system could be valuable in helping to detect abnormal situations more easily.

LEAG provided an extensive dataset for this case study, comprising a ten-year export of the control system archive for power plant Unit Q. The dataset is over 3TB in size and contains signal data from individual measurement points (sensor data like flows, levels, temperatures, and pressures) as well as alarm and event data. The chosen technique to detect abnormal plant conditions makes use of a Dense autoencoder neural network architecture. The functionality of this approach has been validated in previous studies for process industries [23, 24, 25, 26]. The focus of this case study was the development of a dashboard solution to visualize and particularly *explain* the predicted anomalies from this anomaly detection model to the operator. To provide these explanations, technical concepts for several visual *explainers* were developed, that are presented in this paper. The explanations should enable the operator to asses whether the presented situation is indeed an anomaly or not (true or false positive). The proposed dashboard solution is presented in the following section.

4 The Anomaly Explanation Dashboard

Fig. 1 depicts the developed dashboard. The underlying anomaly detection model makes predictions for anomalies for the given plant data, and the dashboard visualizes and explains these predictions to the user. The explainers and feedback components are marked in Fig. 1 by the numbers (D1)-(D8). Two additional elements, (N1) and (N2), have been added after our evaluation workshop, based on suggestions received from interviewees ("N" for "new"). It is worth noting that several elements, namely (D1)-(D4), build on the signal-based input and output data from the *same* underlying Autoencoder anomaly detection model. One explainer-area (D5/D6) supplements these signal-based explanations with alarm log data which is textual data. The textual elements (N1/N2) use text generation techniques to transform some of the numerical information on the dashboard into natural language.

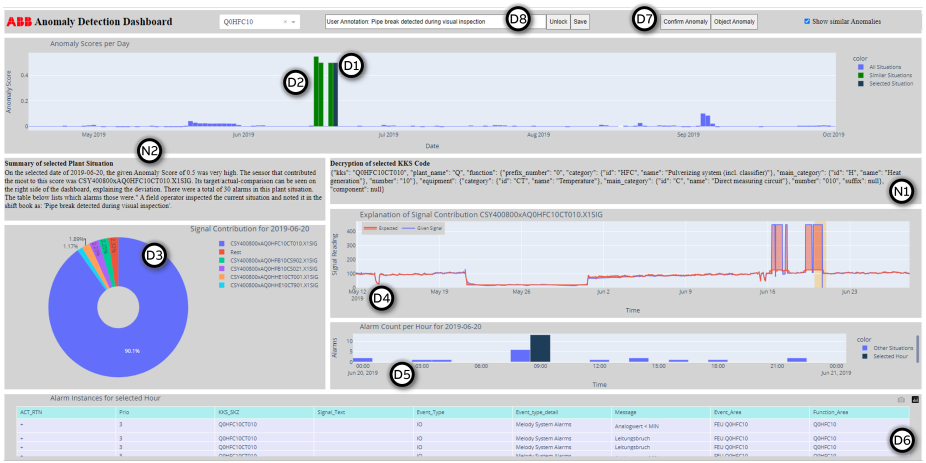

Figure 1: The anomaly detection dashboard, with its explainers and feedback options.

The Explanation Methods

The user starts by selecting a plant area to analyze from a dropdown list on the top left of the dashboard. In the example in Fig. 1, the coal pulverizing system "Q0HFC10" is selected. The histogram below then shows the detected anomalies in this plant area over time. The presence and strength of an anomaly is indicated by an anomaly score. The score is calculated based on the reconstruction error from the underlying Autoencoder model for reconstructing the multivariate time series for the given plant situation, as presented in [24, 25]. To view the explanations for an anomaly, the user can select it from the histogram. The selected situation is then highlighted by dark blue color (D1).

The first explainer is realized in the same histogram: When a plant situation has been selected (D1), the dashboard also searches for any similar plant situations, and highlights them in green (D2). In XAI literature, showing similar situations corresponds to the explanation technique "*case-based explanations*" [11]. Showing similar anomalies can significantly help operators in understanding a given situation, as the root-cause for some of the similar past situations may have already been investigated and thus is known. Say, there was a similar anomaly two weeks ago, and after visual inspection by a field operator it turned out that the problem was due to a clogging issue in a roller mill. Knowing that the present anomaly is again very similar to the one from two weeks ago, provides an indicator that it could be again the same clogging issue.

To find similar anomalies, a separate anomaly clustering model was trained, using the predicted anomalies from the autoencoder as input to the model training. As presented in [24], it is possible to calculate the reconstruction error *per signal* in the multivariate signal data, which provides a kind of "fingerprint" for anomalies. For case-based explanations, these anomaly fingerprints where used as input features to the clustering algorithm, to determine similar anomalies with similar fingerprints. Several clustering algorithms from scikit-learn [27] have been explored for this purpose. First exploratory analysis provided best results for the GaussianMixture clustering. However, more in-depth analysis is needed here which is a further research in this project.

The donut chart in (D3) provides insights into "*feature importance*", another explanation technique [8, 9]. It shows how much different signals contributed to an anomaly. The complexity of equipment in a plant's functional area can provide diverse potential root causes for an anomaly. Being able to pinpoint the anomaly to a specific sensor signal, or a small number of signals, can significantly reduce the search space for field operators. The calculation of the signal contributions makes use again of the signal-wise reconstruction errors. Aggregating the error values for each sensor during the given time window of the anomaly (as presented in [24]) constitutes the total contribution of that sensor to the anomaly that is then displayed in the donut-chart (D3).

When selecting a signal from the donut, the trend chart in (D4) presents two trend lines: A blue trend representing how the signal *actually* progressed in the given (anomalous) situation, and a red trend how it *should have* progressed in the normal case. Being able to directly compare a trend to the expected normal trend can provide additional explanation for the potential anomaly root-cause. This explanation corresponds best to the "*counterfactual explanation*" technique [14] from the state of the art: It presents the model's concepts of the normal state. The calculation of the normal and actual signal trends is again making use of the autoencoder for predicting the anomaly. This model was trained using normal data when the plant operated within normal bounds [24, 25]. When anomalous data is given to the model, a reconstruction error occurs because the model reconstructs this data more closely to how it had learned it. The explainer in (D4) makes use of this fact by visualizing the input signals to the model as actual trends, and the reconstructed output as approximation of the normal trends. In difference to this reconstruction-based approach, the state-of-art counterfactual explanations usually represent the sample with the smallest changes where the model prediction flips (here: from 'anomaly' to 'normal'). Finding the sample with the smallest change usually involves a search or optimization algorithm, whereas using the reconstructed data has the benefit that it is available without additional computation.

The explainers up to this point build on *time series data* from sensor signals. As mentioned, the research team was provided also with the historic alarm logs from the plant. Alarm data contains a lot of textual information, such as an alarm title, and a more detailed alarm description text. Correlating this textual information with an anomaly can help operators to get more context information about the anomaly. In the XAI literature, this explanatory technique is highlighted as "*contextualization*" [12], i.e. to explain by adding domain-specific context information. In the dashboard it is realized in (D5) by showing the amount of alarms that have occurred during an anomalous situation, and (D6) showing the textual alarm details for these alarms.

The aim of the explanations in the dashboard is to enable operators to form their own judgment as to whether those explanations seem plausible if the presented situation is indeed a true anomaly or not (true or false positive). Beyond this operational use, we also aim to provide operators the possibility to give feedback, to be able to retune the models in case predictions were not satisfactory, which is a further research in this project. For this purpose, there is the option in (D7) that allows the user to confirm or object to anomalies. A further feedback option in (D8) allows operators to textually annotate anomalies, e.g. by entering their conclusion. The feedback provided can help model developers to include or exclude data during model retraining, as proposed in [28]. However, textual annotations could also be presented to operators, as additional information in the dashboard to explain similar anomalies in future.

5 User Evaluation

In September 2023, an evaluation workshop was carried out with operators from the LEAG power plant in Boxberg. The workshop aimed to assess how helpful operators found the developed solution in detecting and explaining abnormal plant situations. For the evaluation, we were able to interview five operators during their regular operator training at the LEAG power plant simulator & training center in Lübbenau, Germany

[29]. This came with the benefit of being in an environment similar to their actual control room in Boxberg. Among the respondents, there was a good mix of different levels of work experience, from trainees with only a few months of experience, to operators having been there since the plant was built 40 years ago and seem to know it from the ground up. The respondents' job roles included *control room operators* and *field operators*. A control room operator's primary workplace is in the control room, where the plant is monitored and operated. Field operators are also located in the control room, but are required to go out into the field when a problem is detected, to visually inspect and correct the problem. Younger operators typically start out as field operators and move up to become control room operators as they gain experience. During their training day, we were given a free time slot to present our dashboard to all the operators in the large group. In a second step, we then conducted individual interviews, always with one operator at a time, while the other operators continued with their training. In the interviews, we asked each operator for their feedback on the dashboard using a semi-structured questionnaire. Each interview lasted approximately 15 minutes.

In the first part of the interview, we wanted to get to know the relevance of detecting anomalies today. For all interviewees, the detection of anomalies is an important part of their work. Today, operators detect anomalies by observing the updates of the alarms and message logs, and by constantly monitoring various signal trends. By interpreting anomalous changes in those trends based on their experience, it is possible to detect anomalies and to narrow down their root-causes. However, this analysis requires a lot of experience. Moreover, not all problems can be detected in the control room. There is always the need for a field operator to go out and visually inspect the issue. For example, when standing in front of an engine, one may *hear* that the engine has strange running noises, which is not possible from the control room, as one field operator pointed out.

The frequency of anomalies can vary significantly. There are days, when nothing happens. Operators, however, are in constant anticipation that something *can* occur. Many anomalies can arise, e.g., due to varying coal quality grades coming in from the open pit mines nearby Boxberg. Also plant startup and shutdown phases, e.g. for maintenance or revisions, are a source of anomalies that need to be handled.

In the interviews, all respondents found the presented solution to be a very helpful tool, because it helps to detect and narrow down plant disturbances, which is a crucial part of their work. However, there was also consensus that AI may only alert the operator about anomalies, but should never autonomously intervene when anomalies are detected. The operator should remain in the position to assess the situation themselves, to determine whether it is indeed an anomaly or not, before any actions are taken. That is precisely why it is important to explain anomalies to operators. With regard to a useful integration of the dashboard into the control system, operators could imagine having it on a separate screen, or bringing it up on an existing screen on demand (e.g. in case of an anomaly-alert) because screens are already quite full today. Field operators did not raise a need for an own mobile app, because they also look at the screens together with the control room operators, to discuss anomalies, before going out to the plant to fix them. Here, the presented dashboard was found helpful also by field operators, to discover where to go.

In the second part of the interviews, we asked operators to rate how helpful they found the different dashboard elements (D1)-(D8), for explaining anomalies, and for giving feedback to them. For each dashboard element, we asked interviewees for their rating from "not helpful at all" (0 points) to "absolutely helpful" (4 points). Interviewees could also provide further comments that we noted down in a comment field at each question. The given operator ratings are presented in Table 1, and their corresponding comments, are discussed in the following.

Table 1: Operator’s ratings for the evaluated explainers and feedback options (D1)-(D8).

		Number of responses					average rating
		not helpful at all (0 points)	slightly helpful (1 point)	moderately helpful (2 points)	very helpful (3 points)	absolutely helpful (4 points)	(on the scale 0 - 4 points)
Anomaly detection							
D1	Anomaly scores histogram				4	1	3,2
Anomaly explanation							
D2	Highlight similar situations			1	1	3	3,4
D3	Signal contribution ("donut")					5	4
D4	Actual-vs-normal trend comparison				1	4	3,8
D5, D6	Alarm count & details				3	2	3,4
Operator feedback							
D7	confirm/object anomaly predictions		1	1	2	1	2,6
D8	Annotate anomalies with a text			1	1	3	3,4

Overall, the indication of anomalies with help of the histogram (D1) was rated as very helpful. One operator suggested additionally displaying a threshold line to indicate at which point an anomaly score should be considered very critical, and to alert the operator once this threshold is exceeded, reducing the need for constant monitoring of the dashboard. Furthermore, such a threshold would not have to be fixed, but could be individually adjustable by the user (e.g. through a slider), as some operators may prefer earlier notifications, having a greater need for information than others. Among the various concepts for explaining anomalies, the donut chart in (D3), that analyzes which sensor signal contributed most to an anomaly, was rated to be the most helpful, receiving the highest rating from all respondents. During troubleshooting it would help to determine where to direct the field operator to locate and rectify the issue.

Identifying and showing similar past situations in the dashboard (D2) would help especially when there is already an investigation into why these situations occurred, as field operators do keep a log book about what has been repaired in the past. In those cases the log book entries could suggest a known solution to a similar new problem.

The actual-vs-normal trend comparison in (D4) was mostly rated as "absolutely helpful". Here, interviewees noted that this comparison is not directly displayed to operators in the operator screens today. This means, if there is an anomaly, operators cannot directly see, which upper and lower limits are configured for a sensor. To see how the limits for a signal are set, one would have to check in the so-called control blocks of the process control system, and then it would be possible to compare these normal values with the actual sensor readings in the trend displays, to achieve a kind of actual-vs-normal comparison. However, this process is cumbersome and involves many manual steps.

A similar rationale was provided to us regarding the alarm analysis in (D5/D6). As shown in Table 1, respondents rated this as "very useful" to "extremely useful" for explaining an anomaly. Initially, we were skeptical whether information about alarms would add value to the dashboard, given that all alarms are already displayed in the screens of the control system. However, operators pointed out that there is a benefit to directly see which alarms are related to an anomaly. This speeds up the search process. Otherwise, one would always have to manually search for the relevant alarms in the control system. And speed is crucial when having to make swift decisions.

One suggestion raised in the interviews was to translate the encoded sensor names, shown e.g. in the legend of (D3), into clear text. These codes follows the so-called "KKS" industry standard [30] for identifying different types of equipment in a power plant. A more experienced control room operator is used to read these codes to identify the equipment. However, one cannot expect a young field operator to understand a given KKS name at first glance, when having to go out into the plant. One dashboard extension made after the workshop was therefore the component (N1) to automatically decode the selected sensor names. Due to time constrains, the decoded output is a simple json string, and the visualization of this string can certainly be improved.

Another idea from the discussions was to summarize all the (numerical) explanations for a given anomaly in a short textual paragraph in a natural language format. This is the second addition (N2) made to the dashboard after the workshop interviews, and corresponds to the explanatory technique from XAI literature to "*rephrase model predictions in natural language*" [13]. In the current solution, this is achieved using a template text for the paragraph with space holders, and rules how to textually interpret the analysis from the explainers to fill these place holders. For example, if the predicted anomaly score in (D1) is above a 75 percentile threshold, then this score is translated to be "very high" in the text summary, as shown in the example of (N2) in Fig. 1.

The last part of the interviews concerned the "operator feedback" (D7/D8), i.e. the possibility to provide feedback to an anomaly. This part raised the most concerns from operators. With regard to options (D7) that should allow an operator to confirm anomalies or object to them, interviewees raised the concern about potential "chronic dissenters" who might always click "object". This would lead to many true anomalies being labeled as "false positives", resulting in the unfortunate consequence that those anomaly would not be displayed anymore in future. The option (D8) to annotate anomalies with a short textual description reminded interviewees of their *shift books* they have today, which are also used to log incidents, and looked at again when a similar incident arises. Hence, interviewees rated this feature as helpful, also in our dashboard. However, there might be a challenge when different operators provide varying explanations for an anomaly. Here, the concern was raised that multiple, conflicting annotations could be made for an anomaly, resulting in the need to search for the annotation which are correct, or most useful, whenever a similar anomaly occurs again.

6 Conclusion

In our evaluation interviews with power plant operators, the explainer concepts were predominantly rated as very helpful by the operators to investigate the root causes of the identified anomalies. Most concerns expressed by operators were related to the issue of incorporating user feedback in the tool, e.g., how to deal with contradictory feedback, or how to deal with overly cautious feedback where true anomalies could be labeled as false predictions and therefore not shown anymore in future. It is worthwhile noting that this goes beyond the common objective of XAI, which tries to provide insight into the reasoning of the model. Instead, the combination of model and XAI provides the operator with insight into the process state.

The focus of our research so far has been on developing the explainers to make anomalies understandable, resulting in the dashboard presented in this paper. Our further research will shift more towards the issue of *interactive machine learning (IML)* [28] and also *explanatory interactive machine learning (XIL)* [31], where the given operator feedback in the tool should enable the re-tuning of the anomaly detection models. Here, the expressed concerns raised in our interviews will be valuable input for implementing this feedback component properly. A limitation of the research demonstrator is that it operates *offline* on the exported plant data. To deploy such a solution in a live plant, various technical questions must be addressed, which is not a focus of this research.

Acknowledgments

This work was funded by the German Federal Ministry of Education and Research (BMBF) as per resolution of the German Parliament under the funding code 01IS22030A [6]. We gratefully acknowledge this support. The authors are responsible for the contents of this contribution.

References

[1] Wenjing Lyu and Jin Liu. Artificial intelligence and emerging digital technologies in the energy sector. *Applied energy*, 303:117615, 2021.

[2] Hubert Szczepaniuk and Edyta Karolina Szczepaniuk. Applications of artificial intelligence algorithms in the energy sector. *Energies*, 16(1):347, 2022.

[3] Mengnan Du, Ninghao Liu, and Xia Hu. Techniques for interpretable machine learning. *Communications of the ACM*, 63(1):68–77, 2019.

[4] Amina Adadi and Mohammed Berrada. Peeking inside the black-box: a survey on explainable artificial intelligence (xai). *IEEE access*, 6:52138–52160, 2018.

[5] Feiyu Xu, Hans Uszkoreit, Yangzhou Du, Wei Fan, Dongyan Zhao, and Jun Zhu. Explainable ai: A brief survey on history, research areas, approaches and challenges. In *Natural Language Processing and Chinese Computing: 8th CCF International Conference, NLPCC 2019, Dunhuang, China, October 9–14, 2019, Proceedings, Part II 8*, pages 563–574. Springer, 2019.

[6] EU-funded research project EXPLAIN - web page. `https://explain-project.eu/`. Accessed: 2023-10-23.

[7] Nicolas Pfeuffer, Lorenz Baum, Wolfgang Stammer, Benjamin M Abdel-Karim, Patrick Schramowski, Andreas M Bucher, Christian Hügel, Gernot Rohde, Kristian Kersting, and Oliver Hinz. Explanatory interactive machine learning. *Business & Information Systems Engineering*, pages 1–25, 2023.

[8] Marco Tulio Ribeiro, Sameer Singh, and Carlos Guestrin. " why should i trust you?" explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, pages 1135–1144, 2016.

[9] Scott M Lundberg and Su-In Lee. A unified approach to interpreting model predictions. *Advances in neural information processing systems*, 30, 2017.

[10] Oscar Li, Hao Liu, Chaofan Chen, and Cynthia Rudin. Deep learning for case-based reasoning through prototypes: A neural network that explains its predictions. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 32, 2018.

[11] Rich Caruana, Hooshang Kangarloo, John David Dionisio, Usha Sinha, and David Johnson. Case-based explanation of non-case-based learning methods. In *Proceedings of the AMIA Symposium*, page 212. American Medical Informatics Association, 1999.

[12] Freddy Lecue. On the role of knowledge graphs in explainable ai. *Semantic Web*, 11(1):41–51, 2020.

[13] Erik Cambria, Lorenzo Malandri, Fabio Mercorio, Mario Mezzanzanica, and Navid Nobani. A survey on xai and natural language explanations. *Information Processing & Management*, 60(1):103111, 2023.

[14] Ruth MJ Byrne. Counterfactuals in explainable artificial intelligence (xai): Evidence from human reasoning. In *IJCAI*, pages 6276–6282, 2019.

[15] Udo Schlegel, Hiba Arnout, Mennatallah El-Assady, Daniela Oelke, and Daniel A Keim. Towards a rigorous evaluation of xai methods on time series. In *2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW)*, pages 4197–4201. IEEE, 2019.

[16] Ioana Giurgiu and Anika Schumann. Explainable failure predictions with rnn classifiers based on time series data. *arXiv preprint arXiv:1901.08554*, 2019.

[17] Sarthak Manas Tripathy, Ashish Chouhan, Marcel Dix, Arzam Kotriwala, Benjamin Klöpper, and Ajinkya Prabhune. Explaining anomalies in industrial multivariate time-series data with the help of explainable ai. In *2022 IEEE International Conference on Big Data and Smart Computing (BigComp)*, pages 226–233. IEEE, 2022.

[18] Thomas Rojat, Raphaël Puget, David Filliat, Javier Del Ser, Rodolphe Gelin, and Natalia Díaz-Rodríguez. Explainable artificial intelligence (xai) on timeseries data: A survey. *arXiv preprint arXiv:2104.00950*, 2021.

[19] Emanuel Metzenthin. Lime-for-time. `https://github.com/emanuel-metzenthin/Lime-For-Time`, 2017.

[20] Felix Mujkanovic, Vanja Doskoč, Martin Schirneck, Patrick Schäfer, and Tobias Friedrich. timexplain–a framework for explaining the predictions of time series classifiers. *arXiv preprint arXiv:2007.07606*, 2020.

[21] Arzam Kotriwala, Benjamin Klöpper, Marcel Dix, Gayathri Gopalakrishnan, Dawid Ziobro, and Andreas Potschka. Xai for operations in the process industry-applications, theses, and research directions. In *AAAI Spring Symposium: Combining Machine Learning with Knowledge Engineering*, pages 1–12, 2021.

[22] Official website of the LEAG Power Plant in Boxberg, Germany. `https://www.leag.de/de/geschaeftsfelder/kraftwerke/kraftwerk-boxberg/`. Accessed: 2023-10-23.

[23] Marcel Dix, Benjamin Kloepper, Jean-Christophe Blanchon, and Elise Thorud. A formula for accelerating autonomous anomaly detection. *ABB Review*, 2:2021, 2021.

[24] Marcel Dix, Ashish Chouhan, Srishti Ganguly, Sripada Pradhan, Devika Saraswat, Surbhi Agrawal, and Ajinkya Prabhune. Anomaly detection in the time-series data of industrial plants using neural network architectures. In *2021 IEEE Seventh International Conference on Big Data Computing Service and Applications (BigDataService)*, pages 222–228. IEEE, 2021.

[25] Marcel Dix and Srishti Ganguly. Ai-based anomaly detection of asset failures in industrial process plants. *atp magazin*, 64(5):60–67, 2022.

[26] Arzam Kotriwala, Ruomu Tan, Pablo Rodriguez, Marcel Dix, Benedikt Schmidt, and Anne Lene Rømuld. Plant operator support using industrial artificial intelligence. *atp magazin*, 65(10):70–76, 2023.

[27] Scikit-Learn - Overview of clustering methods. `https://scikit-learn.org/stable/modules/clustering.html`. Accessed: 2023-10-23.

[28] Saleema Amershi, Maya Cakmak, William Bradley Knox, and Todd Kulesza. Power to the people: The role of humans in interactive machine learning. *Ai Magazine*, 35(4):105–120, 2014.

[29] LEAG power plant simulator in Luebbenau, Germany. `https://www.konferenzcenter-spreewald.de/de/kraftwerkssimulator`. Accessed: 2023-10-23.

[30] VGBE-Standard: Kraftwerk-Kennzeichensystem (KKS). `https://www.vgb.org/shop/kks.html/`. Accessed: 2023-10-23.

[31] Stefano Teso and Kristian Kersting. Explanatory interactive machine learning. In *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*, pages 239–245, 2019.