2024 Tenth Indian Control Conference (ICC)
9 - 11 December, 2024
IISER Bhopal, India

Machine learning based explainable fault detection of vacuum pumps in PV solar cell manufacturing

Ajay Narayanan
Reliance Industries Limited
Bangalore, India
ajay1.narayanan@ril.com

Dr. Kedar Kulkarni
Reliance Industries Limited
Bangalore, India
kedar5.kulkarni@ril.com

Fery Riswan
REC Solar
Singapore
fery.riswan@recgroup.com

Muhammad Azzrul Bin Salehan
REC Solar
Singapore
azzrul.muhd@recgroup.com

***Abstract*—Pump failures in industrial settings can lead to significant operational disruptions and financial losses. This paper presents a data-driven approach for explainable fault detection in a vacuum pump operating within a chemical vapor deposition (CVD) unit at a Photovoltaic (PV) solar cell manufacturing facility. A comprehensive dataset was collected and several machine learning based classification models were evaluated for their ability to predict pump failures based on the pump sensor data. All models demonstrated exceptional performance, achieving F1-scores ranging from 0.96 to 0.98. Furthermore, explainability analysis on the best model revealed that temperature-related features were the most influential in predicting faults. Additionally, counterfactual analysis provided actionable recommendations to adjust operational parameters in order to mitigate future failures. These findings highlight the potential of machine learning for proactive maintenance and optimization in industrial settings.**

***Index Terms*—PV solar cell manufacturing, fault detection and diagnosis, machine learning, explainable AI, counterfactual analysis**

I. Introduction

Industrial processes heavily rely on critical equipment for their seamless operation. Unforeseen failures in such machinery can trigger substantial financial losses due to production downtime, repair costs, and potential safety hazards [1]. In the context of solar PV cell manufacturing, vacuum pumps play a pivotal role in maintaining the optimal vacuum conditions within chemical vapor deposition (CVD) chambers, essential for producing high-quality and efficient solar cells [2]. Any malfunction in these pumps can disrupt production and compromise product quality, making fault detection and diagnosis (FDD) a critical concern for manufacturers [3].

Traditionally, fault detection strategies for industrial equipment involve manual inspection and expert knowledge, that can often lead to unexpected downtime and inefficient resource allocation. To address this, data-driven approaches, particularly those leveraging machine learning (ML), have gained prominence in recent years [4]. However, the complexity and 'black-box' nature of many ML models can make their adoption in industrial settings challenging [5]. Plant operators and engineers often require transparent and interpretable explanations for model predictions to build trust and help effective decision-making. This has led to the rise of explainable AI (XAI) techniques [6], which aim to provide human-understandable insights into the factors influencing model predictions.

979-8-3503-1416-8/23/$31.00 ©2023 IEEE

Related work: Numerous studies have demonstrated the effectiveness of machine learning (ML) for FDD in various types of pumps, including centrifugal and reciprocating pumps [4]. Kumar et al. [7] employed Support Vector Machines (SVMs) for fault detection in centrifugal pumps, achieving high accuracy in identifying various fault conditions. In another study, Chang et al. [8] used a combination of statistical features and ML models, including Random Forest and XGBoost, to predict failures in centrifugal pumps, achieving impressive results in terms of accuracy and early detection. Sayed et al. [9] explored the use of Artificial Neural Networks (ANNs) for fault detection in centrifugal pumps, demonstrating their ability to detect incipient faults with high accuracy. Haslhofer et al. [10] developed an anomaly detection model based on Long Short-Term Memory (LSTM) networks to detect faults in dry vacuum pumps used in semiconductor manufacturing.

Despite the successes of ML-based FDD, the interpretability of these models remains a challenge. Explainable AI (XAI) techniques [6] have thus emerged as a critical component in enhancing the transparency and trust in ML-based FDD systems. Several generic XAI approaches exist in literature that make ML models explainable or human-interpretable. For example, SHAP (SHapley Additive exPlanations) proposed by Antwarg et al. [11] can be used for comprehensive feature importance analysis. At the same time, counterfactual analysis offers another promising avenue to enhance the interpretability and actionability of ML models. Mothilal et al. [12] proposed DiCE (Diverse Counterfactual Explanations), a framework for generating diverse and actionable counterfactual explanations, providing actionable recommendations to lead the system to a different, potentially desired outcome. Bhakte et al. [13] explored

the use of counterfactual explanations for fault diagnosis in industrial systems, demonstrating their potential for providing actionable insights to operators. While existing research has explored ML-based FDD and XAI in various contexts, their application to CVD vacuum pumps within solar PV cell manufacturing remains relatively underexplored.

This study addresses this gap by presenting a systematic approach for explainable fault detection in a CVD vacuum pump at a PV solar cell manufacturing facility. By harnessing machine learning and XAI, the proposed approach not only predicts potential faults but also indicates the underlying factors contributing to these anomalies. This dual approach enables both accurate fault detection and actionable insights for plant operators to prevent future faults. The contributions of this work are twofold. First, it demonstrates the effectiveness of ML and XAI in the context of industrial fault detection, providing a real-world example of their successful integration. Second, the study generates valuable insights on the specific factors influencing vacuum pump failures, offering practical recommendations for optimizing pump operation and preventing costly downtime. By bridging the gap between data-driven predictions and actionable insights, this work aims to pave the way for more effective and efficient maintenance strategies in the PV solar manufacturing industry and beyond.

II. Methodology

This section details the proposed systematic approach for explainable fault detection in a CVD vacuum pump at a PV solar cell manufacturing facility. The proposed approach has four key stages: (1) data collection and preprocessing (2) model development and evaluation (3) explainability analysis (4) counterfactual analysis. Each of these stages is explained in the following subsections.

A. Data collection and preprocessing

Operational data for a CVD vacuum pump at a PV solar cell manufacturing facility was extracted from on-board sensors. The acquired time-series data encompassed various parameters relevant to the pump's performance, as detailed in Table I. Overall, 60 days of sensor data was collected for this study. The data collection spanned periods of normal operation, identified as a 'good period' (day 1 to day 30 and day 49 to day 60), and the period showing abnormal operation, identified as the 'bad period' (day 31 to day 48). Given the high quality of the data collected, minimal feature engineering was required, allowing for a focus on the raw sensor values as the primary predictors of pump health.

For further analysis, the cleaned dataset was divided into training and test sets. The training set comprised data from day 1 to day 38, while the test set encompassed the period from day 39 to day 60. To ensure the proper functioning of the classification models and enable them to learn from both 'good period' and 'bad period' data, a portion of the 'bad period' data (specifically, from day 31 to day 38) was included in the training set. This partitioning strategy provides the models with a balanced representation of 'good period' and 'bad period' while reserving a portion of the 'bad period' data (day 39 to day 48) for evaluation in the test set.

B. Model development and evaluation

A diverse array of machine learning *classification* models was considered for predicting the pump's operational state (good/bad). This included tree-based models (RandomForest [14], XGBoost [15]), boosting models (AdaBoost [16]), linear models (Perceptron [17], SGDClassifier [18]), and graph-based semi-supervised models (Label Spreading [19], Label Propagation [20]). Ensemble techniques like CalibratedClassifierCV [21] and kernel-based methods like SVC [22] were also explored. The performance of each trained model was evaluated on the unseen test set using various metrics, including precision, recall and the F1-score. The results were compared to identify the best predictive model in terms of accuracy.

C. Explainability analysis

Explainability analysis was then conducted to gain insights into the decision-making process of the fault detection model. This involved the computation of SHAP (SHapley Additive exPlanations) [11] values for each feature, quantifying their contribution to the model's predictions. By examining the SHAP values, the relative importance of each feature in determining the classification of 'good period' and 'bad period' was determined. This analysis highlighted the key factors influencing the pump's behavior and provided a deeper understanding of the underlying mechanisms leading to potential failures.

D. Counterfactual analysis

Counterfactual analysis is a technique that explores "what-if" scenarios by perturbing input features to providing actionable recommendations to lead the system to a different, potentially desired outcome [23].In this study, the Diverse Counterfactual Explanations (DiCE) framework [12] , with its genetic search algorithm, was used to generate *counterfactual instances* that represent feasible and realistic modifications to the pump's operational parameters, such that it pushes the operational state of the vacuum pump from a 'bad period' (which is detected by the ML model) to a 'good period'. These counterfactual explanations provided valuable insights into how specific adjustments could potentially prevent or mitigate pump failures, which is extremely useful for plant operators.

TABLE I
DESCRIPTION OF VARIABLES ACQUIRED FROM CVD VACUUM PUMP SENSORS

Variable name	**Variable description**
MB Motor Current (I_{MB})	Current flowing through the motor connected to the blower (MB)
DP Motor Current(I_{DP})	Current flowing through the motor connected to the dry pump (DP)
Motor Speed (N)	Rotational speed of the motor
TC Temperature (T_{TC})	Temperature reading from the thermo couple(TC) unit

III. RESULTS AND DISCUSSION

A. Exploratory data analysis

Exploratory data analysis (EDA) reveals distinct patterns in the operational parameters of the vacuum pump over time, as depicted in Fig 1 . The figure comprises four subplots, each displaying the time series of a specific parameter. The numerical values of the parameters depicted in this figure have been normalized to lie between 0 and 1. The data points are color-coded to distinguish between normal operation ('good periods', green dots) and the period showing faulty pump operation ('bad periods', red dots). As is evident from the figure, T_{TC} experiences a sharp spike during the 'bad period' which suggests that this parameter will be crucial for fault detection. I_{MB} remains relatively stable during normal operation but exhibits fluctuations and an upward trend in the 'bad period'. In contrast, the motor speed (N) and I_{DP} shows consistent behavior throughout, with only minor deviations. These observations highlight the complex interplay of various operational parameters in the pump's performance.

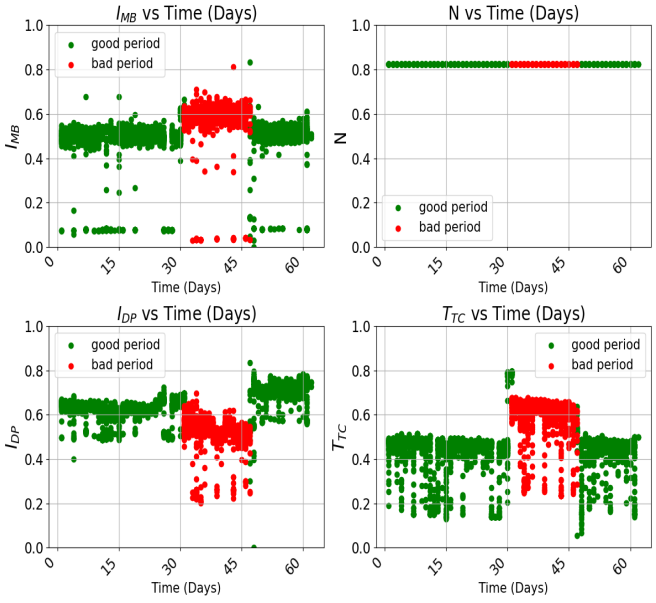


Fig. 1. Exploratory data analysis

B. Fault detection results

As discussed in section 2.2, several models were trained and tested on an Apple M1 Pro with 16 GB RAM, yielding execution times of less than 2 seconds. The evaluation of various classification models for fault detection in the CVD vacuum pump, as presented in the Table II , reveals consistently exceptional performance across all models. Notably, the F1-scores of all models ranged from 0.96 to 0.98, highlighting the models' strong ability to accurately identify both good and bad operational periods based on sensor data. The minimal variation in execution times as well as their F1-scores, indicates that any of these models can be effectively deployed. In this work, we choose the RandomForest model since it has the highest F1-sore on the test set and, being a tree-based model, it is lends itself easily to interpretability.

Fig 2 illustrates the predictions of the RandomForest model on the test set for a single parameter I_{MB}, where the shaded grey regions denote the model-predicted 'bad periods'. As is evident from the figure, the model demonstrates exceptional performance with a precision of 98.6%, indicating a low false positive rate of 2.4%. Furthermore, it achieves a high recall of 99.2%, capturing almost all actual 'bad periods'. The overall F1-score of 98% underscores the model's excellent balance between precision and recall.

C. Explainability Analysis

To enhance the interpretability of the RandomForest model, explainability analysis was conducted, as visualized by the SHAP bee swarm summary plot shown in Fig 3 . This figure displays an information-dense summary of how the operational parameters of the pump impact the model's prediction of pump failures. T_{TC} emerged as the

TABLE II
FAULT DETECTION MODEL COMPARISON RESULTS

Models	Precision	Recall	F1-score
RandomForest	0.986	0.992	0.98
AdaBoost	0.976	0.99	0.976
Perceptron	0.97	0.97	0.97
LabelSpreading	0.968	0.973	0.97
LabelPropogation	0.974	0.968	0.971
XGBClassifier	0.965	0.979	0.969
SVC	0.973	0.964	0.965
SGDClassifier	0.968	0.973	0.964

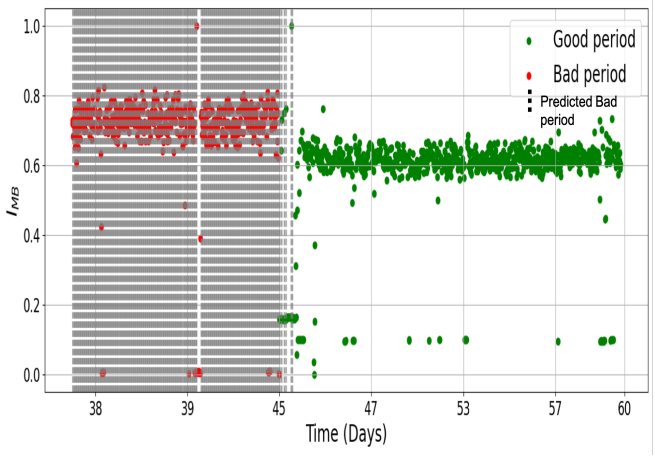


Fig. 2. Fault detection results by RandomForest model

most influential feature, with higher temperatures strongly associated with a higher probability of fault occurrence. The current parameters, I_{MB} and I_{DP}, also contribute positively to the failure prediction, though to a lesser extent. In contrast, the motor speed (N) appears to have a negligible impact on the model's output, as indicated by its low SHAP values. This analysis highlights that the temperature-related feature, T_{TC}, was the most influential in predicting faults.

D. Counterfactual Analysis

For counterfactual analysis, some random instances in the test set, predicted as the 'bad period' by the RandomForest model, were selected. Choosing these random instances as the *original points*, we used the DiCE framework to generate

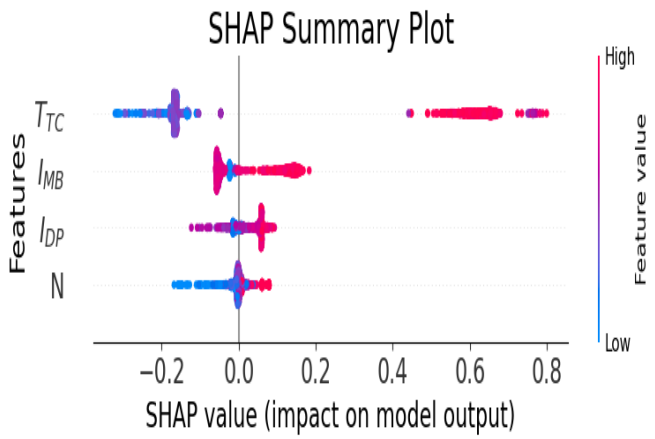


Fig. 3. SHAP beeswarm summary plot

a set of *counterfactual instances* for each of these *original points*, as discussed in section 2.4.

Table III-D presents counterfactual recommendations for three distinct *original points* (Points 1-3) all of which were predicted as a 'bad period' by the RandomForest model. For each *original point*, two *counterfactual instances* were generated as shown. In general, the *counterfactual instances* recommend a decrease in the T_{TC} and I_{MB} parameters to alter the operational state of the pump from a 'bad period' to a 'good period'.

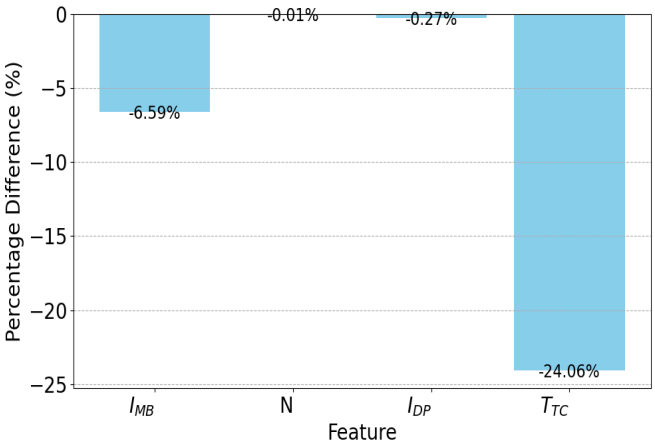


Fig. 4. Average % difference between counterfactual instances and original point

Fig 4 shows the average percentage difference between the *counterfactual instances* and all three original points across the four features . T_{TC} exhibits the most significant difference (-24.06%), indicating that a decrease in this parameter is the most influential factor in altering the model's prediction from 'bad period' to 'good period'. I_{MB} also shows a

TABLE III
COUNTERFACTUAL ANALYSIS RESULTS

Point	Feature	Original Feature Values	Counterfactual Instance 1	Counterfactual Instance 2
Point 1	I_{MB}	0.6	0.5	0.47
	I_{DP}	0.56	0.55	0.54
	T_{TC}	0.8	0.53	0.45
	N	0.81	0.80	0.81
Point 2	I_{MB}	0.65	0.52	0.4
	I_{DP}	0.56	0.55	0.54
	T_{TC}	0.83	0.55	0.49
	N	0.81	0.80	0.81
Point 3	I_{MB}	0.62	0.52	0.49
	I_{DP}	0.56	0.55	0.54
	T_{TC}	0.82	0.51	0.48
	N	0.81	0.80	0.81

considerable difference (-6.59%), suggesting its potential role in mitigating pump failures. In contrast, I_{DP} and N exhibit very small differences (-0.27% and -0.01%, respectively). These results align with the discussion provided in section 3.3, in that T_{TC} and I_{MB} have a much greater influence on model predictions than I_{DP} and N.

To better visualize the *counterfactuals instances* generated by the DiCE framework, we visualize the feature space, as shown in Fig 5, for by I_{MB} and T_{TC} – the two parameters that showed the most significant change. In this figure, the 'good' and 'bad' period instances are placed in regions colored according to their classification by the RandomForest model: 'good period' (purple) and 'bad period' (yellow). For illustration, we choose one of the 3 original points (shown in red). The counterfactual instance, representing the nearest 'good period' instance achievable through minimal feature manipulation, is marked by the green triangle. The figure illustrates how, based on counterfactual analysis, reducing T_{TC} and I_{MB} can help the CVD vacuum pump transition from a 'bad period' to a 'good period'.

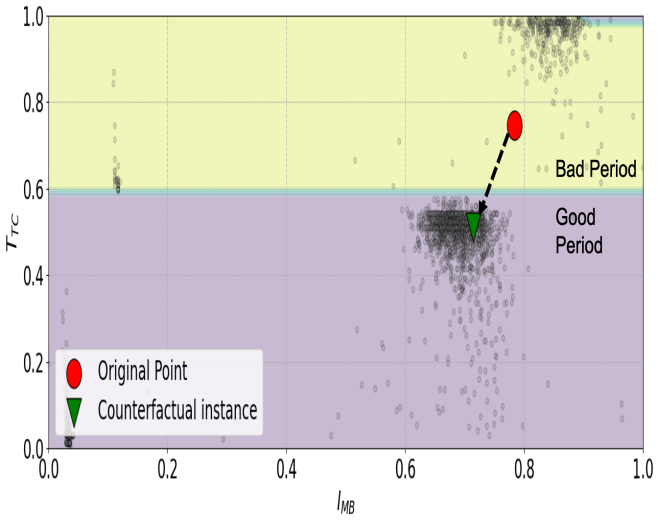


Fig. 5. Counterfactual analysis results in the space of features I_{MB} and T_{TC}.

IV. CONCLUSION

This study demonstrates the effectiveness of the proposed systematic approach for explainable fault detection in a CVD vacuum pump that is critical to PV solar cell manufacturing. By leveraging sensor data and training various ML-based classification models, the proposed approach accurately identified 'bad periods', achieving F1-scores ranging from 96% to 98%. Explainability analysis using SHAP values revealed T_{TC} as the most influential factor in predicting failures, a finding consistent with the physical understanding of pump malfunctions. Additionally, counterfactual analysis using the DiCE framework generated actionable recommendations to adjust operational parameters to mitigate future failures.

REFERENCES

[1] "What is Machine Downtime? Definition, Examples and Impacts — output.industries," [Accessed 29-06-2024].

[2] "Photovoltaic solar cell fabrication with vacuum chamber in Energy Transition — blog.vinci-technologies.com," [Accessed 29-06-2024].

[3] Y.-J. Park, S.-K. S. Fan, and C.-Y. Hsu, "A review on fault detection and process diagnostics in industrial processes," *Processes*, vol. 8, no. 9, p. 1123, 2020.

[4] A. Angelopoulos, E. T. Michailidis, N. Nomikos, P. Trakadas, A. Hatziefremidis, S. Voliotis, and T. Zahariadis, "Tackling faults in the industry 4.0 era—a survey of machine-learning solutions and key aspects," *Sensors*, vol. 20, no. 1, p. 109, 2019.

[5] L. E. Lwakatare, A. Raj, I. Crnkovic, J. Bosch, and H. H. Olsson, "Large-scale machine learning systems in real-world industrial settings: A review of challenges and solutions," *Information and software technology*, vol. 127, p. 106368, 2020.
[6] F. Xu, H. Uszkoreit, Y. Du, W. Fan, D. Zhao, and J. 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*. Springer, 2019, pp. 563–574.
[7] A. Kumar and R. Kumar, "Time-frequency analysis and support vector machine in automatic detection of defect from vibration signal of centrifugal pump," *Measurement*, vol. 108, pp. 119–133, 2017.
[8] K. Chang and S. H. Park, "Random forest-based multi-faults classification modeling and analysis for intelligent centrifugal pump system," *Journal of Mechanical Science and Technology*, vol. 38, no. 1, pp. 11–20, 2024.
[9] E. Sayed, A. A. Abdelsamee, and N. M. Ghazaly, "Classification model using neural network for centrifugal pump fault detection," *International Journal*, vol. 13, no. 4, 2020.
[10] B. Haslhofer and C. Heistracher, "Machine learning methods for analysis and prognosis of outages in reactive ion etching processes in the semiconductor industry," 2021.
[11] L. Antwarg, R. M. Miller, B. Shapira, and L. Rokach, "Explaining anomalies detected by autoencoders using shapley additive explanations," *Expert systems with applications*, vol. 186, p. 115736, 2021.
[12] R. K. Mothilal, A. Sharma, and C. Tan, "Explaining machine learning classifiers through diverse counterfactual explanations," in *Proceedings of the 2020 conference on fairness, accountability, and transparency*, 2020, pp. 607–617.
[13] A. Bhakte and R. Srinivasan, "Potential for counterfactual explanations to support digitalized plant operations," in *Computer Aided Chemical Engineering*. Elsevier, 2024, vol. 53, pp. 3073–3078.
[14] G. Biau and E. Scornet, "A random forest guided tour," *Test*, vol. 25, pp. 197–227, 2016.
[15] T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," in *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, 2016, pp. 785–794.
[16] R. E. Schapire, "Explaining adaboost," in *Empirical inference: festschrift in honor of vladimir N. Vapnik*. Springer, 2013, pp. 37–52.
[17] S. I. Gallant *et al.*, "Perceptron-based learning algorithms," *IEEE Transactions on neural networks*, vol. 1, no. 2, pp. 179–191, 1990.
[18] R. M. Gower, N. Loizou, X. Qian, A. Sailanbayev, E. Shulgin, and P. Richtárik, "Sgd: General analysis and improved rates," in *International conference on machine learning*. PMLR, 2019, pp. 5200–5209.
[19] "LabelSpreading — scikit-learn.org," [Accessed 29-06-2024].
[20] "LabelPropagation — scikit-learn.org," [Accessed 29-06-2024].
[21] "CalibratedClassifierCV — scikit-learn.org," [Accessed 29-06-2024].
[22] S. Suthaharan and S. Suthaharan, "Support vector machine," *Machine learning models and algorithms for big data classification: thinking with examples for effective learning*, pp. 207–235, 2016.
[23] S. Verma, J. Dickerson, and K. Hines, "Counterfactual explanations for machine learning: A review," *arXiv preprint arXiv:2010.10596*, vol. 2, p. 1, 2020.