https://doi.org/10.62836/jcmea.v3i1.030108

Global Science Publishing

JCMEA

Journal of Computational Methods in Engineering Applications

https://ojs.sgsci.org/journals/jcmea

Article

CORAL-based Domain Adaptation Algorithm for Improving the Applicability of Machine Learning Models in Detecting Motor Bearing Failures

Guojun Zhang [1], Tong Zhou [2,*] and Yiqun Cai [3]

[1] *Quadrant International Inc., San Diego, CA92121, USA*

[2] *Air China Cargo Co., Ltd., Beijing101318, China*

[3] *University of Florida, Gainesville, FL32608, USA*

Abstract: Motor bearings are essential components in various industrial and transportation systems, vital for minimizing friction and enhancing machinery longevity. Failures in these bearings can lead to extensive machine downtime and significant repair costs, thereby emphasizing the need for effective predictive maintenance strategies. This paper focuses on leveraging advancements in Machine Learning (ML) and Artificial Intelligence (AI) to preemptively identify and rectify potential bearing failures, transitioning from traditional periodic maintenance to more efficient, condition-based approaches. We introduce a novel domain adaptation technique using Correlation Alignment (CORAL) to improve the accuracy of fault predictions across different operational settings. This method effectively minimizes the statistical disparities between training and operational data, enhancing the adaptability and effectiveness of predictive models. The results indicate that models equipped with domain adaptation outperform traditional models, particularly in their ability to generalize across diverse environments, thereby supporting more reliable and efficient predictive maintenance practices. This research contributes to the ongoing evolution of maintenance strategies in industrial settings, highlighting the potential of AI to transform traditional practices by reducing unexpected downtime and optimizing maintenance schedules.

Keywords: motor beaing failtures; machine learning; domain adaptation

1. Introduction

Motor bearings are important components in the machinery of various industrial and transportation systems [1,2], playing a critical role in improving smooth operations and extending the lifespan of these systems. Their primary function is to reduce the friction between moving parts, ensuring that machinery operates at optimal levels of efficiency. However, bearing failures can lead to significant machinery breakdowns, necessitating costly repairs and causing extensive downtime, which can adversely affect practical industrial applications such as chemical processing and transportation [3,4]. Therefore, the ability to predict and address potential bearing faults is highly advantageous for maintaining operational reliability and efficiency.

Motor bearing failures can occur due to a range of issues, including wear and tear, insufficient lubrication, contamination by foreign substances, or mechanical misalignment. Each of these factors can degrade bearing

Received: 18 October 2023; Accepted: 3 November 2023.

* Corresponding: Tong Zhou (zhoutong@airchinacargo.com)

performance and, if left unchecked, lead to failures that disrupt operations and pose safety risks. The consequences of such failures are not minor; they can severely impact the production processes and safety protocols of facilities that depend heavily on motor-driven machinery. Thus, predictive maintenance, which focuses on predicting failures before they occur, has become an indispensable strategy in industrial operations. This proactive approach replaces traditional periodic maintenance schedules with maintenance actions that are precisely timed based on the actual condition of the equipment.

The advent of Machine Learning (ML) and Artificial Intelligence (AI) [5–7] has revolutionized in many domains. For example, Xiong et al. proposed an ensemble model of attention mechanism-based DCGAN and autoencoder for effective noised OCR classification [8]. Traditional methods, which often relieve simple statistical tools and manual inspections used in many studies [9 – 11], are increasingly being replaced by sophisticated AI-driven models. These models are capable of analyzing large datasets collected from sensors monitoring machine conditions, such as temperature, vibration, and acoustics. By using historical data, these models learn to identify patterns and anomalies that precede mechanical failures. This capability allows them to provide early warnings of potential bearing faults, enabling timely interventions that can prevent costly breakdowns and enhance machine longevity.

Despite the advancements in AI technologies [12–14], applying these models across different machines or operational environments presents significant challenges. Specifically, the variability in data due to different operating conditions, machine types, or environmental factors can lead to substantial performance declines when models trained in one domain are applied to another. This issue of data variability and model performance under different conditions is known as the problem of distribution shifts. To overcome this, domain adaptation techniques are employed, which are designed to enable AI models to adapt from one domain (the source domain) to another (the target domain).

Domain adaptation is particularly critical in the context of motor bearing fault prediction [15–17]. It can address the issue where models developed and trained under specific sets of conditions fail to generalize well across other conditions that exhibit different data characteristics [18]. This technique involves adjusting the model or its training process so that it can effectively handle data from different domains, thereby improving its accuracy and reliability across diverse operational settings.

This article introduces a novel approach to domain adaptation shown in Figure 1 for enhancing the prediction accuracy of motor bearing failures. The first step in our proposed method involves segmenting the dataset into different clusters using the k-means clustering algorithm. This segmentation is based on the premise that each cluster represents distinct data distributions. Following this, a source domain is selected—one that is rich in labeled data and representative of typical operational conditions. The next phase involves aligning the data distributions between the source domain and other clusters using a technique known as Correlation Alignment (CORAL). CORAL minimizes the statistical differences between the source and target distributions, effectively making the model more adaptable and robust to changes in data characteristics across different domains. Once the domain adaptation is accomplished, the refined datasets serve as training material for several machine learning models, including K-Nearest Neighbors (KNN), Random Forest, and Decision Trees. These models are selected for their ability to handle classification tasks effectively and are applied to predict the likelihood of bearing failures. The performance of these models, both with and without domain adaptation, is evaluated to assess the impact of our adaptation technique. We anticipate that models trained with domain-adapted data will exhibit superior performance, demonstrating enhanced generalizability and robustness compared to models trained on non-adapted data.

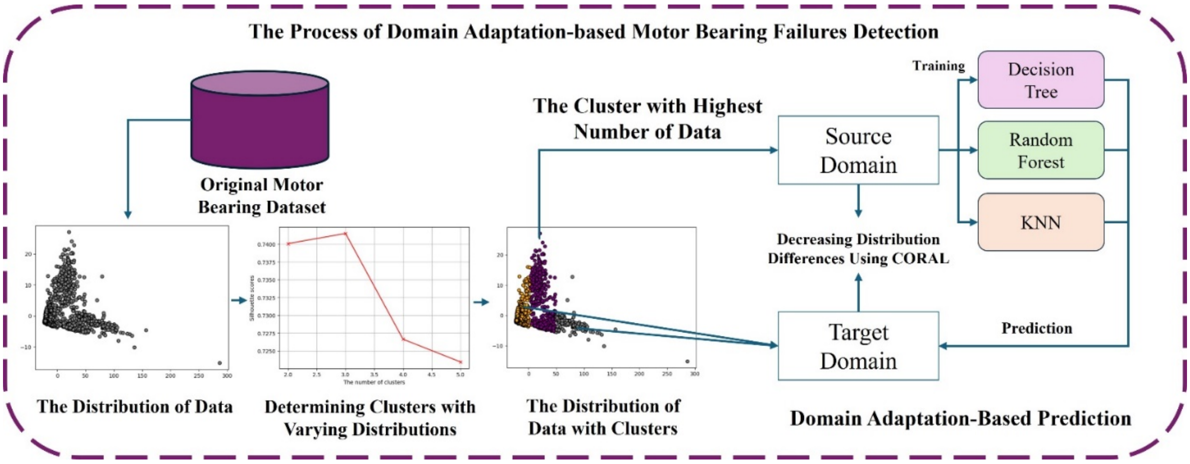


Figure 1. The process of the proposed motor bearing failures detection method using CORAL-based domain adaptation.

2. Literature Review

2.1. Bearing Failures Detection

Recent literature on bearing fault diagnosis demonstrates a growing focus on the application of advanced machine learning techniques used in many studies [19–21] to improve detection accuracy. For instance, Wang et al. explored the use of wavelet packet transform combined with sparse representation theory to enhance fault feature extraction in rolling element bearings [22]. Similarly, Han et al. developed a multiscale convolutional neural network that leverages data augmentation to improve the robustness of bearing fault diagnosis systems [23]. Another critical area of research is the adaptation of models to varying conditions. Zhang et al. proposed a sparse decomposition-based method for bearing fault diagnosis that adapts to different operational conditions, addressing the challenge of varying data distributions [24]. Additionally, Sharma et al. emphasized the use of nonlinear dynamic analysis using Higuchi's fractal dimension to detect defects in rolling element bearings, highlighting the importance of sophisticated analytical techniques in the diagnosis process [25]. These studies underscore the integration of complex data processing techniques and machine learning models to address the dynamic challenges in bearing fault diagnosis.

2.2. Domain Adaptation

Despite progress in AI technologies [26,27], deploying these models across diverse machines or operational contexts remains a considerable challenge. Domain adaptation is then considered to solve these issues. Pan et al. categorize transfer learning based on domain and task variations into three types: inductive, transductive, and unsupervised [28]. Inductive involves different tasks, possibly with the same domain, requiring some labeled target data. Transductive has the same tasks across domains but uses labeled source data and some target domain's unlabeled data to learn its distribution. Unsupervised transfer has different tasks and domains, with no labeled data, relying on entirely unlabeled sets.

In domain adaptation, methods are split into shallow and deep techniques. Shallow strategies, such as instance-based and feature-based adaptations [29 – 31], use metrics like maximum mean discrepancy [32], Wasserstein metric, correlation alignment [33], and others to minimize domain differences. Deep adaptation employs neural networks with architectures like convolutional, autoencoder, or adversarial, integrating distance metrics to address discrepancies in feature representation across domains.

3. Method

3.1. Dataset Descrption and Preprocessing

The dataset used in this study, provided by Case Western Reserve University through its Bearing Data Center, aims to support the application of machine learning in predictive maintenance of industrial machinery. It is particularly focused on motor bearing fault detection and classification using telemetry from mechanical components under specific test conditions. The motor specifications are 2 HP motor with defects introduced via EDM machining at points with diameters of 0.007 inches, 0.014 inches, and 0.021 inches. The Measurement tools are Torque transducer, dynamometer, and control electronics. The dataset contains time-series data segmented into 2048-point intervals (approximately 0.04 s at the 48 kHz sampling rate). Nine features are calculated for these segments, including maximum, minimum, mean, standard deviation, RMS, skewness, kurtosis, crest factor, and form factor, which are essential for identifying potential faults. The dataset serves to diagnose faults specifically in three parts of the bearing: the ball, inner race, and outer race. The distribution of features and the target variable are plotted in Figures 2 and Figure 3. It can be found that each category in the target variable has the same number that equals to 230.

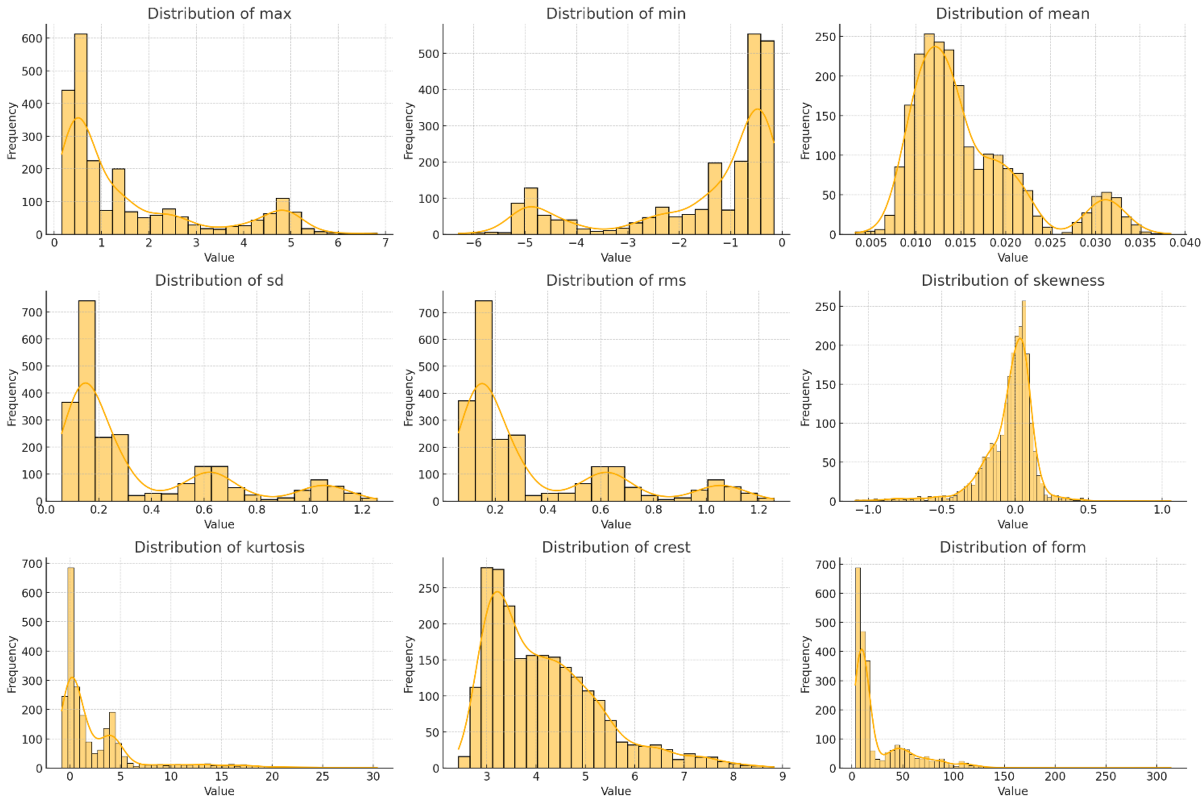


Figure 2. The distribution of features.

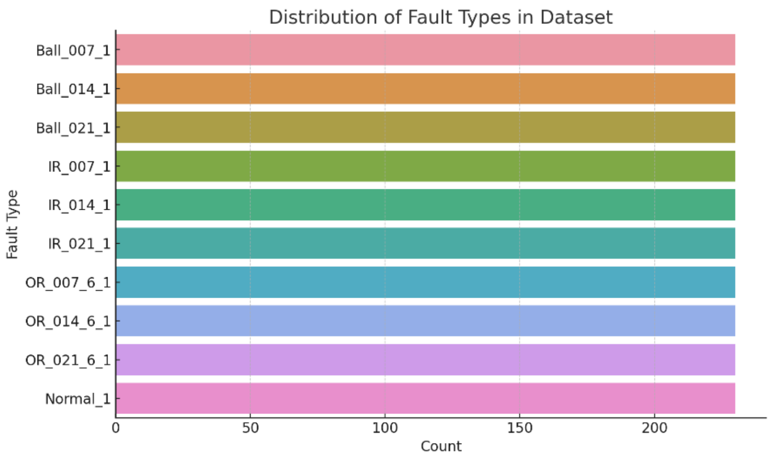


Figure 3. The distribution of target variable.

During the data preprocessing phase of our research, we initially converted both categorical features and the target variable into numerical formats via one-hot encoding. Subsequently, we standardized the feature values using min-max normalization to ensure uniformity in scale across all variables, thereby aiding in the effectiveness and speed of our model training process. To address the need for domain adaptation, we segmented the motor bearing data into distinct clusters using the K-means clustering method, a popular unsupervised technique that organizes data into ‘k’ groups by reducing the distances between data points and their corresponding cluster centers. To make sure the ideal number of clusters (k), we utilized the Silhouette score, a measure that assesses the appropriateness of data points within their own clusters relative to other clusters, with higher scores indicating more precise clustering. Our results indicated that the most effective clustering occurred with k = 3. We identified the largest cluster, visually represented in purple, as the primary source domain, while the smaller clusters, shown in orange and grey, were designated as the target domains. The distribution of the original data, the progression of the Silhouette score, and the post-clustering PCA distribution are depicted in Figures 4–6 of our paper, respectively.

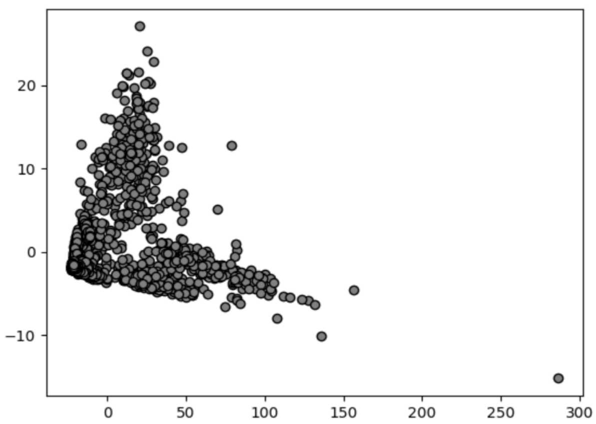


Figure 4. The PCA distribution of the original data.

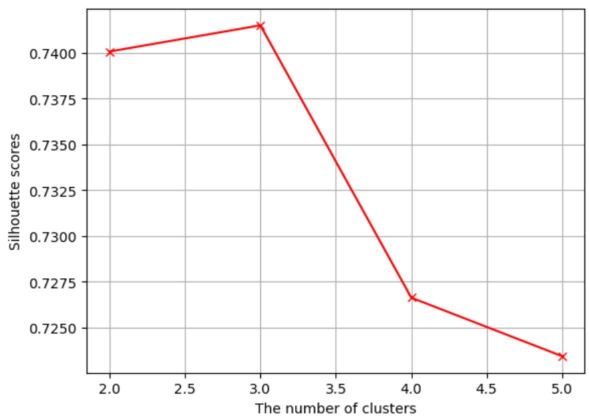


Figure 5. The curve of Silhouette scores used for determining k value.

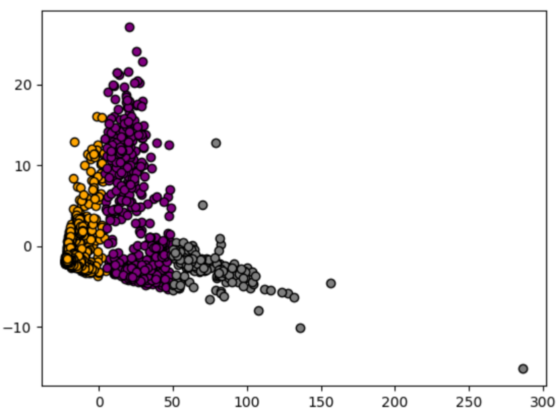


Figure 6. The PCA distribution of the data with segmented clusters.

3.2. Domain Adaptation-Based Machine Learning Models

3.2.1. Decision Tree

Decision tree [34, 35] was first considered in this study due to its excellent prediction performance and interpretability. A decision tree is a versatile machine learning algorithm used for both classification and regression tasks. At its core, it involves splitting data into branches to make predictions, forming a tree-like structure of decisions. The tree is built from a root node and expands by branching off into possible outcomes based on the features of the data. The decision tree starts at the root node, which holds the entire dataset. At each node, the tree asks a question about one of the features based on set criteria (such as Gini impurity, entropy, or variance reduction), and branches off into nodes according to the answers to these questions. This process continues recursively, creating a flowchart-like structure. Each branch represents an outcome, and each leaf node represents a final decision or prediction. One of the most significant advantages of decision trees is their ease of interpretation and visualization. They mimic human decision-making closer than other algorithms, making them intuitive and straightforward to explain even to non-technical stakeholders. Additionally, decision trees require relatively little data preparation. Unlike many other algorithms, they don't require feature scaling or centering at all [36 – 38]. Moreover, decision trees serve as the foundational building blocks for more advanced ensemble methods such as Random Forests and Gradient Boosting Machines, which combine multiple decision trees to improve performance and overcome some of the overfitting issues associated with single decision trees.

3.2.2. Random Forest

A Random Forest [36, 37] is an ensemble machine learning technique that builds upon the simplicity of decision trees and enhances their performance and accuracy. It consists of a large number of individual decision trees that operate as an ensemble. Each individual tree in the Random Forest spits out a class prediction and the class with the most votes becomes the model's prediction. The fundamental concept behind Random Forest is a simple but powerful one—the wisdom of crowds. In data science speak, the reason that the Random Forest model works so well is that a large number of relatively uncorrelated models (trees) operating as a committee will outperform any of the individual constituent models. Random Forest applies the general technique of bootstrap aggregating, or bagging, to tree learners. Each tree in the ensemble is built from a sample drawn with replacement (i. e., a bootstrap sample) from the training set. Additionally, when splitting a node during the construction of the tree, the split that is chosen is no longer the best split among all features. Instead, the split that is picked is the best split among a random subset of the features. As a result, this strategy, called the random subspace method, ensures that the trees are de-correlated. When it comes to making predictions, the Random Forest aggregates the decisions of multiple trees to decide on the final output. For classification tasks, this typically means a majority voting mechanism, whereas for regression tasks, it usually involves averaging the predictions from each tree. One of the biggest advantages of Random Forest is its robustness. It is one of the most accurate learning algorithms available, which provides a good indicator for the robustness of the method. For many data sets, it produces a highly accurate classifier. The Random Forest algorithm is also known for its ability to limit overfitting. Since it takes the ensemble of relatively uncorrelated models, it won't overfit the data, unlike a single decision tree might.

3.2.3. K-Nearest Neighbors

The K-Nearest Neighbors (KNN) [38] algorithm is a simple, easy-to-implement supervised machine learning algorithm that can be used to solve both classification and regression problems. It's a type of instance-based learning, or lazy learning, where the function is only approximated locally and all computation is deferred until classification. KNN works by finding the closest data points in the training set to the new point and predicts the output based on these nearest neighbors. The 'K' in KNN is a parameter that refers to the number of nearest neighbors to include in the majority voting process. The distance between points is usually calculated using Euclidean distance, although other metrics such as Manhattan or Minkowski can also be used depending on the context. For classification tasks, KNN assigns the new data point the class most common among its k nearest neighbors. In regression tasks, it assigns the property value based on the average of the values of its k nearest neighbors. KNN is incredibly simple to understand and implement and has shown remarkable effectiveness in cases with nonlinear data. Unlike most other algorithms, KNN doesn't require training time because the model structure isn't determined until a query is made to the system. This makes it uniquely flexible. In addition, KNN can easily handle cases with more than two classes without any increase in complexity or decrease in performance.

3.2.4. Domain Adaptation

CORAL, short for Correlation Alignment [33], is a machine learning technique that addresses the challenge of domain adaptation, which is particularly vital when training and testing data come from different distributions. This technique aims to align the second-order statistics (covariances) of source and target domains to reduce the domain shift, thereby enhancing the performance of a model when applied to the target domain.

Domain adaptation through CORAL minimizes the difference between the feature distributions of the source and target datasets. The central idea is that even if two datasets represent similar tasks but come from different sources (e. g., different sensors, environments, or times), their data distributions might differ significantly. CORAL aligns these distributions by adjusting the covariances of the source domain to match those of the target domain. This adjustment makes the features from both domains more comparable, allowing models trained on the source domain to perform better on the target domain without extensive retraining.

The process involves adjusting the data from the source domain so that its covariance aligns with the target domain. This is achieved by computing a linear transformation that modifies the source features. By applying this transformation, the variance of the features in the source domain is made similar to that in the target domain, thereby aligning the multi-dimensional data distributions. The transformed source data can then be used to train models that are more likely to generalize well on the target data.

CORAL is computationally efficient compared to other domain adaptation methods. It does not require complex optimization or extensive hyper-parameter tuning, making it easy to implement and integrate into existing workflows. In addition, The technique is versatile and can be applied across a variety of tasks where domain adaptation is required, including computer vision, natural language processing, and speech recognition. Whether the shifts between domains are due to different acquisition conditions, varying operational settings, or temporal changes, CORAL has been effective in bridging the gap, thereby ensuring model robustness.

4. Results and Discussion

4.1. The Performance of the Decision Tree

The experimental results presented showcase the performance of a decision tree classifier across two distinct target domains, with a focus on both the original and domain-adapted prediction accuracies. The performance metrics considered include Accuracy, F1 Score, Precision, and Recall.

From Tables 1 and 2, it's evident that the decision tree classifier performs variably across the two domains. For Target Domain-1 (orange), the original performance metrics are notably low with an accuracy of 0.02, F1 score of 0.03, precision of 0.10, and recall of 0.05. This trend suggests significant challenges in adapting the decision tree model to this particular domain using the original training methodology. Conversely, Target Domain-2 (grey) displays a better initial performance with an accuracy of 0.37, F1 score of 0.19, precision of 0.40, and recall of 0.27. These results indicate a relatively higher compatibility or easier adaptability of the decision tree model within this domain, likely due to features or data distributions that are more in line with the model's training base.

Table 1. Original prediction performance using the decision tree.

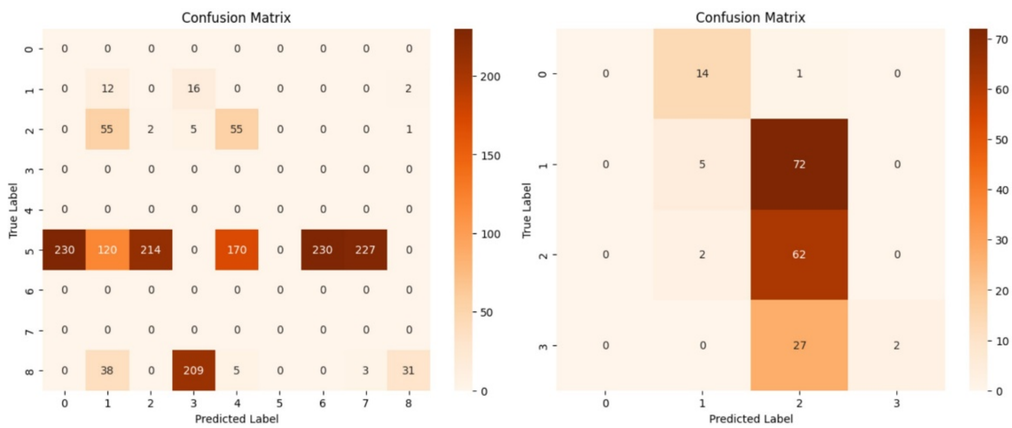
Metrics	Target Domain-1 (Orange)	Target Domain-2 (Grey)
Accuracy	0.02	0.37
F1 score	0.03	0.19
Preicision	0.10	0.40
Recall	0.05	0.27

Table 2. Domain adaptation-based prediction performance using the decision tree.

Metrics	Target Domain-1 (Orange)	Target Domain-2 (Grey)
Accuracy	0.03	0.61
F1 score	0.06	0.56
Preicision	0.04	0.64
Recall	0.10	0.59

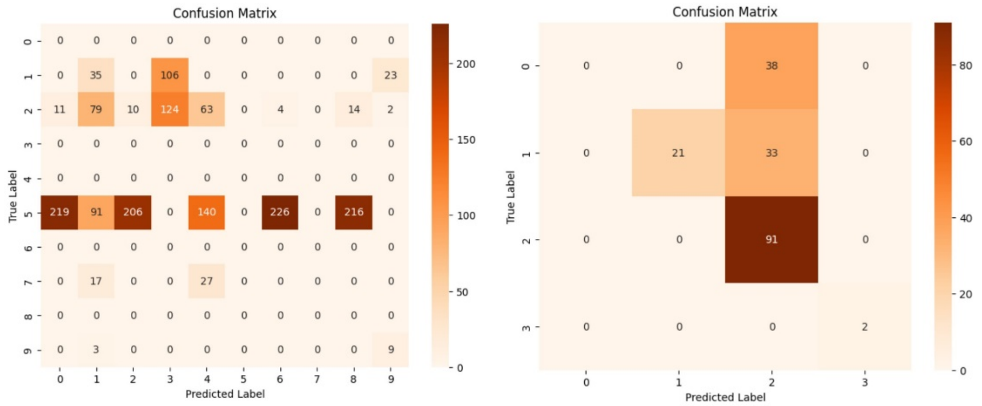
The domain adaptation approach aims to enhance the model's ability to generalize across different datasets by aligning the statistical properties of the source and target domain data. As seen in Table 2, this strategy yields varied outcomes: (1) For Target Domain-1, domain adaptation slightly improves all metrics (accuracy to 0.03, F1 score to 0.06, precision to 0.04, recall to 0.10). Although the improvements are modest, they signify some positive shift in model adaptability due to domain alignment techniques. (2) Target Domain-2 shows a substantial increase in all performance metrics with domain adaptation, with accuracy jumping to 0.61, F1 score to 0.56, precision to 0.64, and recall to 0.59. These improvements underscore the effectiveness of domain adaptation in environments where the original model performance was already somewhat aligned with the domain characteristics.

The confusion matrices (Figures 7 and 8) provide deeper insight into the classifier's performance. For Target Domain-1, the high concentration of misclassifications (especially for labels 5 and 6) before and after adaptation indicates persistent challenges in correctly predicting certain classes. For Target Domain-2, the adaptation leads to a clearer diagonal distribution in the confusion matrix, implying better true positive rates across most classes. Figure 9 also provides the comparative analysis of original and adapted models. It visually underscores the enhancements brought about by domain adaptation, particularly in the context of accuracy and recall metrics, where the adapted model for Target Domain-2 shows substantial gains.



(a) Target domain-orange **(b) Target domain-grey**

Figure 7. Original confusion matrix using the decision tree.



(a) Target domain-orange **(b) Target domain-grey**

Figure 8. Domain-adaptation-based confusion matrix using the decision tree.

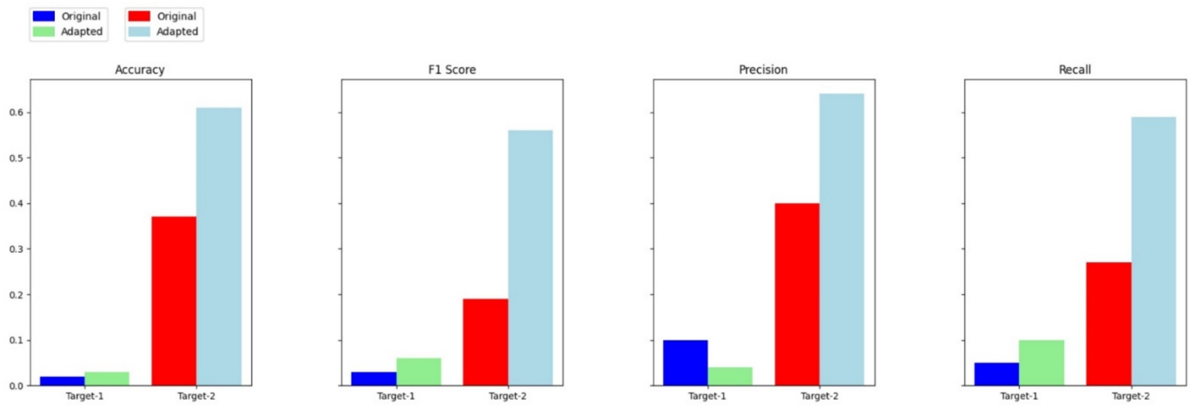


Figure 9. Comparison of Original and Domain Adaptation Performance Across Metrics using the decision tree.

4.2. The Performance of the Random Forest

The original performance of the Random Forest in Target Domain-1 (orange) was markedly low shown in Table 3, with an accuracy of only 0.02 and a F1 score of 0.02. This suggests that the model, without adaptation, struggled considerably to generalize the predictive patterns from the training data to this particular domain. Precision was slightly higher at 0.10, while recall was at a minimal 0.01, indicating a poor ability to correctly identify true positives within the dataset. Conversely, in Target Domain-2 (grey), the Random Forest achieved much higher initial metrics, with an accuracy of 0.45 and a F1 score of 0.34. Precision in this domain was notably robust at 0.79, complemented by a recall of 0.43. These figures suggest a comparatively better initial fit for the model's predictions in this domain, likely due to more favorable alignment between the model's training and the characteristics of the grey domain's data. The confusion matrices from Figure 10 further elucidate the model's performance nuances. For Target Domain-1, the matrix shows widespread misclassifications across multiple classes, with particularly high false negatives and false positives in classes that dominate the dataset, reflecting the low recall and precision values. For Target Domain-2, while there is still some degree of misclassification, the matrix shows a higher concentration of correct predictions in the primary classes, indicating a better model fit.

Table 3. Original prediction performance using the random forest.

Metrics	Target Domain-1 (Orange)	Target Domain-2 (Grey)
Accuracy	0.02	0.45
F1 score	0.02	0.34
Preicision	0.10	0.79

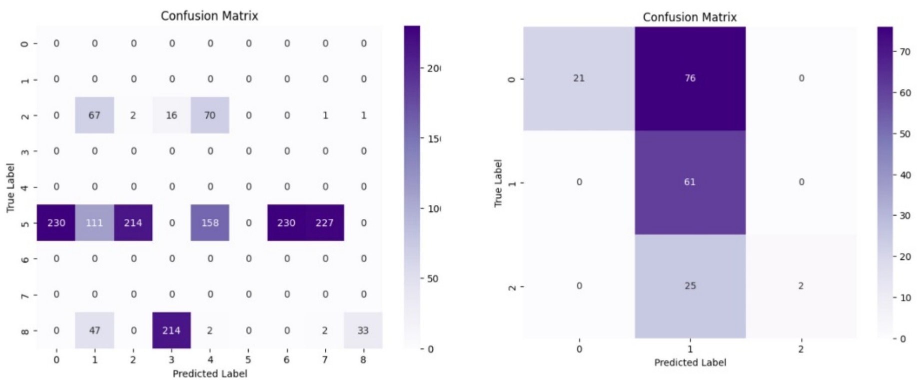
Cont.

Metrics	Target Domain-1 (Orange)	Target Domain-2 (Grey)
Recall	0.01	0.43

After domain adaptation shown in Table 4, both domains exhibited improved metrics, albeit with varying degrees of enhancement. In Target Domain-1, accuracy improved to 0.10 and F1 score to 0.05, which, while still low, represent a doubling of the model's initial performance metrics. Precision and recall saw modest improvements. In Target Domain-2, domain adaptation pushed the accuracy up to 0.57 and the F1 score to 0.50, with precision and recall also experiencing significant increases. The confusion matrices in Figure 11 after domain adaptation reveal a denser concentration of correct predictions, especially in Target Domain-2, where class alignments are noticeably better than in the pre-adaptation state.

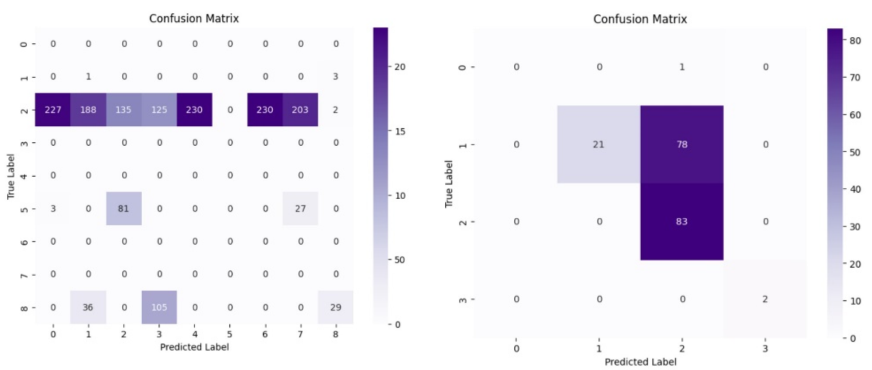
Table 4. Domain adaptation-based prediction performance using the random forest.

Metrics	Target Domain-1 (Orange)	Target Domain-2 (Grey)
Accuracy	0.10	0.57
F1 score	0.05	0.50
Preicision	0.16	0.62
Recall	0.05	0.55



(a) Target domain-orange **(b) Target domain-grey**

Figure 10. Original confusion matrix using the random forest.



(a) Target domain-orange **(b) Target domain-grey**

Figure 11. Domain-adaptation-based confusion matrix using the random forest.

The bar charts in Figure 12 illustrate the stark contrast between the original and adapted model performances. While the adaptation has significantly enhanced performance in both domains, the improvements are more pronounced in Target Domain-2.

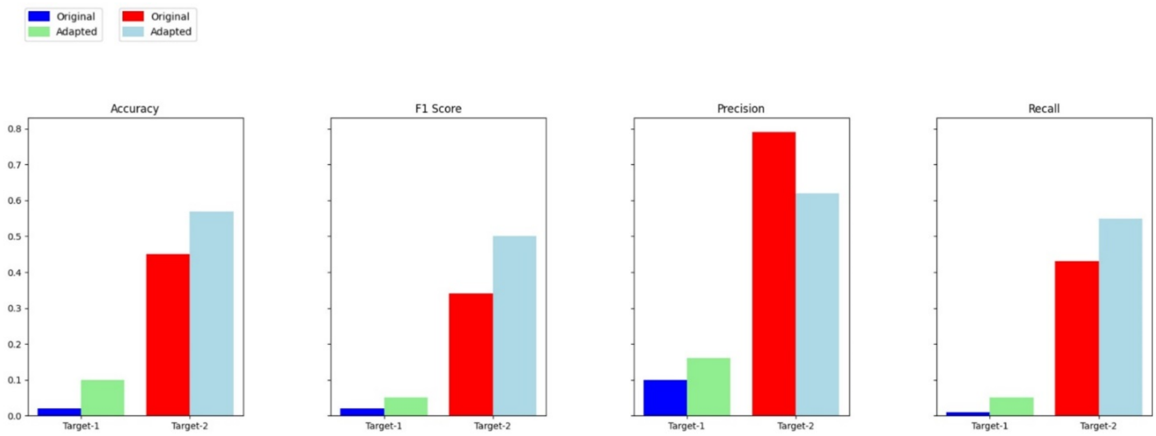


Figure 12. Comparison of Original and Domain Adaptation Performance Across Metrics using the random forest.

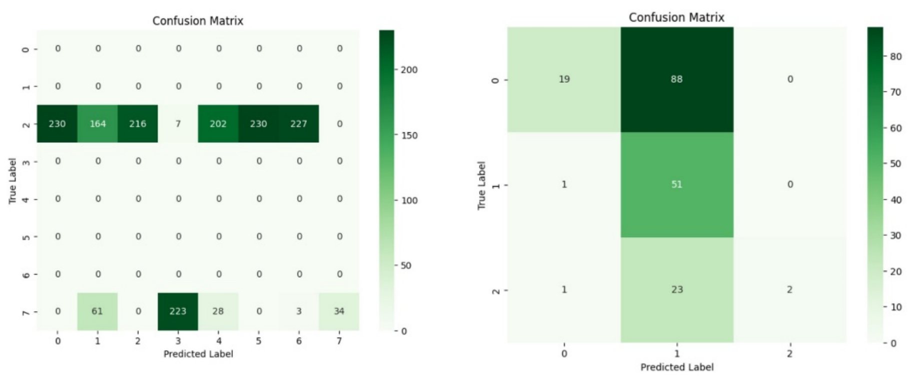
4.3. The performance of the KNN

For Target Domain-1 (orange), the KNN algorithm struggled significantly with an accuracy of just 0.15, an F1 score of 0.05, precision at 0.25, and recall at a low 0.03 shown in Table 5. These metrics suggest that KNN was poorly aligned with the characteristics of this target domain initially. In contrast, Target Domain-2 (grey) exhibited substantially better results with an accuracy of 0.38, F1 score of 0.30, precision at 0.73, and recall at 0.41, indicating a more favorable initial response to this domain's features.

Table 5. Original prediction performance using the KNN.

Metrics	Target Domain-1 (Orange)	Target Domain-2 (Grey)
Accuracy	0.15	0.38
F1 score	0.05	0.30
Preicision	0.25	0.73
Recall	0.03	0.41

The confusion matrices (Figure 13) highlight the challenges in Target Domain-1 with widespread misclassifications across nearly all classes, reflecting the low performance metrics. For Target Domain-2, the matrix shows fewer off-diagonal entries, particularly for one major class, indicating a relatively better alignment of KNN's classification boundaries with the data distribution of this domain.



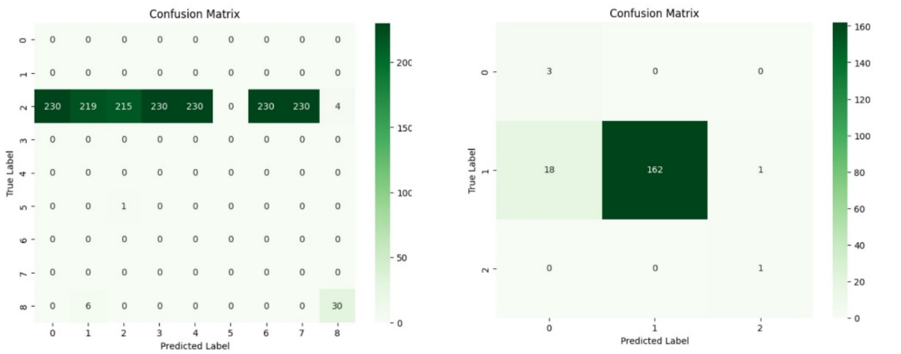
(a) Target domain-orange **(b) Target domain-grey**

Figure 13. Original confusion matrix using the KNN.

Post domain adaptation, the performance in Target Domain-1 saw some improvement; accuracy remained at 0.15 but F1 score increased to 0.12, precision rose to 0.20, and recall increased slightly to 0.10 shown in Table 6. Target Domain-2 experienced significant enhancements post-adaptation, with accuracy soaring to 0.89, F1 score reaching 0.62, precision at 0.54, and recall at an impressive 0.96. The adapted confusion matrices (Figure 14) for Target Domain-1 still show considerable misclassifications but with a slight improvement in the correct predictions for some classes. In Target Domain-2, the adaptation process seems to have effectively resolved the misclassification issues, with a high concentration of correct predictions and very few errors, as seen in the nearly uniform color intensity across the major class.

Table 6. Domain adaptation-based prediction performance using the KNN.

Metrics	Target domain-1 (orange)	Target domain-2 (grey)
Accuracy	0.15	0.89
F1 score	0.12	0.62
Preicision	0.20	0.54
Recall	0.10	0.96



(a) Target domain-orange **(b) Target domain-grey**

Figure 14. Domain-adaptation-based confusion matrix using the KNN.

The bar charts in Figure 15 dramatically illustrate the improvements. While the adaptation effect in Target Domain-1 is visible but modest, in Target Domain-2, there is a remarkable enhancement across all metrics, particularly in recall and F1 score.

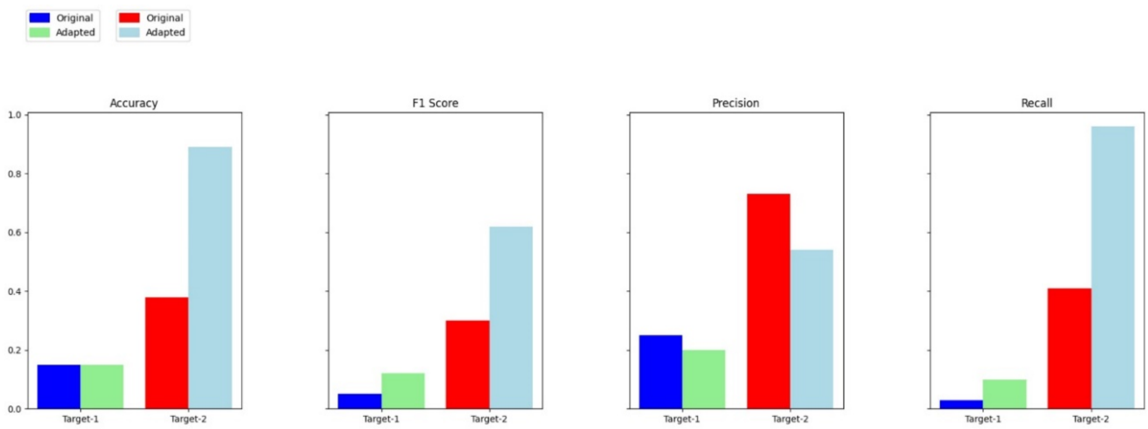


Figure 15. Comparison of Original and Domain Adaptation Performance Across Metrics using the KNN.

The data presented in Table 7 highlights the effectiveness of domain adaptation techniques, particularly the application of CORAL in reducing covariance differences between the source and target domains. Initially, the covariance differences were significantly high, with values of approximately 60.59 in Target Domain-1 and

126.06 in Target Domain-2. These large differences suggest a substantial disparity in data distributions between the source and each target domain, potentially leading to poor model performance when algorithms trained on the source domain are applied directly to the target domains. After applying CORAL, the covariance differences dramatically decreased to near-zero values (0.000005 for Target Domain-1 and 0.000012 for Target Domain-2). This dramatic reduction indicates that the feature distributions of the target domains have been effectively aligned with those of the source domain, facilitating a much smoother transfer of learning models and likely enhancing their predictive accuracy on the target data. Such alignment is crucial for achieving robust performance across different domains and showcases the power of domain adaptation techniques in overcoming challenges posed by dataset variability.

Table 7. The comparison of previous and new covariance difference.

Previous Covariance Difference-Target Domain-1	New Covariance Difference-Target Domain-1	Previous Covariance Difference-Target Domain-2	New Covariance Difference-Target Domain-2
60.586413	0.000005	126.058967	0.000012

4.4. Discussion

Across different models and target domains, the performances of the decision tree, random forest, and KNN algorithms show significant variability. Generally, the random forest and KNN models outperform the decision tree, especially after domain adaptation, indicating the robustness of ensemble and nearest neighbors techniques over simpler decision trees in handling complex data patterns and domain shifts. In the original settings, the decision tree and KNN models struggle particularly in Target Domain-1, likely due to its challenging characteristics that do not align well with these models' assumptions or their capacity to handle noise and feature diversity. Post domain adaptation, there is a noticeable improvement in all models' performance metrics, with particularly striking enhancements in accuracy and recall for the random forest and KNN in Target Domain-2. This suggests that domain adaptation techniques, such as aligning distributions or feature re-scaling, are highly effective for these models.

However, despite these improvements, there remain gaps in performance, especially in Target Domain-1 across all models, which might be attributed to intrinsic data complexities or insufficient model complexity. This highlights a need for further research into more sophisticated domain adaptation strategies or potentially exploring hybrid models that combine the strengths of various learning algorithms to enhance adaptability and predictive accuracy. The current study also suggests exploring deeper into the characteristics of each domain might yield insights that could lead to better-targeted adaptations, enhancing model performance uniformly across domains. But the domain adaptation still can provide excellent performance, suggesting its potential in many domains such as electronics, transportation [39–42] and communications [43–45].

5. Conclusion

The exploration into predictive maintenance for motor bearings using ML and AI demonstrates significant potential to enhance operational reliability and efficiency. The application of domain adaptation techniques, particularly CORAL, effectively reduces discrepancies in data distributions across different operational domains, thereby improving the generalizability and accuracy of predictive models. However, the research also uncovers persistent challenges in model performance across varied conditions, underscoring the need for further enhancement of domain adaptation methods. Future work should focus on refining these techniques and exploring hybrid models that integrate various AI approaches to better handle the complexities of real-world applications [46–54]. This study sets the groundwork for more resilient predictive maintenance strategies that can adapt to diverse and changing industrial environments, thus mitigating the risk of costly machinery failures [57–67].

Funding

Not applicable.

Author Contributions

Conceptualization, G. Z., T. Z. and Y. C.; writing—original draft preparation, G. Z., T. Z. and Y. C.; writing—review and editing, G.Z., T.Z. and Y.C.; All of the authors read and agreed to the published the final manuscript.

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Data Availability Statement

Not applicable.

Conflicts of Interest

The authors declare no conflict of interest.

References

1 Wittek E, Kriese M, Tischmacher H, *et al*. Capacitances and Lubricant Film Thicknesses of Motor Bearings under Different Operating Conditions. In Proceedings of the XIX International Conference on Electrical Machines-ICEM2010, Rome, Italy, 6–8 September 2010; pp. 1–6.

2 Devaney MJ, Eren L. Detecting Motor Bearing Faults. *IEEE Instrumentation & Measurement Magazine* 2004; **7**(**4**): 30–50.

3 Dai W. Evaluation and Improvement of Carrying Capacity of a Traffic System. *Innovations in Applied Engineering and Technology* 2022; 1–9.

4 Yu L, Li J, Cheng S, *et al.* Secure Continuous Aggregation in Wireless Sensor Networks. *IEEE Transactions on Parallel and Distributed Systems* 2013; **25**(**3**): 762–774.

5 Xiong S, Chen X, Zhang H. Deep Learning-Based Multifunctional End-to-End Model for Optical Character Classification and Denoising. *Journal of Computational Methods in Engineering Applications* 2023; 1–13.

6 Wenjun D, Fatahizadeh M, Touchaei HG, et al. Application of Six Neural Network-Based Solutions on Bearing Capacity of Shallow Footing on Double-Layer Soils. *Steel and Composite Structures*2023; **49**(**2**): 231–244.

7 Du M, Liu N, Hu X. Techniques for Interpretable Machine Learning. *Communications of the ACM* 2019; **63**(**1**): 68–77.

8 Xiong S, Zhang H, Wang M. Ensemble Model of Attention Mechanism-Based DCGAN and Autoencoder for Noised OCR Classification. *Journal of Electronic & Information Systems* 2022; **4**(**1**): 33–41.

9 Ren P, Zhao Z, Yang Q. Exploring the Path of Transformation and Development for Study Abroad Consultancy Firms in China. *Journal of Computational Methods in Engineering Applications* 2023; 1–12.

10 Lei J. Green Supply Chain Management Optimization Based on Chemical Industrial Clusters. *Innovations in Applied Engineering and Technology* 2022; 1–17.

11 Li C, Tang Y. The Factors of Brand Reputation in Chinese Luxury Fashion Brands. *Journal of Integrated Social Sciences and Humanities* 2023; 1–14.

12 Giger ML. Machine learning in Medical Imaging. *Journal of the American College of Radiology* 2018; **15**(**3**): 512–520.

13 Kononenko I. Machine Learning for Medical Diagnosis: History, State of the Art and Perspective. *Artificial Intelligence in Medicine*2001; **23**(**1**): 89–109.

14 Richards BA, Lillicrap TP, Beaudoin P, *et al*. A Deep Learning Framework for Neuroscience. *Nature Neuroscience*2019; **22**(**11**): 1761–1770.

15 Farahani A, Voghoei S, Rasheed K, *et al.* A Brief Review of Domain Adaptation. In *Advances in Data Science and Information Engineering, Proceedings of the ICDATA 2020 and IKE 2020, Las Vegas, NV, USA, 27–30 July 2020*; Springer: Cham, Switzerland, 2021; pp. 877–894.

16 Ben-David S, Blitzer J, Crammer K, *et al*. Analysis of Representations for Domain Adaptation. In *Advances in Neural Information Processing Systems*; MIT Press: Cambridge, MA, 2006; Volume 19.

17 You K, Long M, Cao Z, *et al*. Universal Domain Adaptation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition2019, Long Beach, CA, USA, 15–20 June 2019; pp. 2720–2729.

18 Ganin Y, Lempitsky V. Unsupervised Domain Adaptation by Backpropagation. In Proceedings of the 32nd International Conference on Machine Learning, PMLR, Lille, France, 7–9 July 2015; pp. 1180–1189.

19 Xiong S, Zhang H, Wang M, *et al.* Distributed Data Parallel Acceleration-Based Generative Adversarial Network for Fingerprint Generation. *Innovations in Applied Engineering and Technology* 2022; 1–12.

20 Murdoch WJ, Singh C, Kumbier K, *et al.* Definitions, Methods, and Applications in Interpretable Machine Learning. *Proceedings of the National Academy of Sciences*2 019; **116**(**44**): 22071–22080.

21 Hao Y, Chen Z, Jin J, *et al.* Joint Operation Planning of Drivers and Trucks for Semi-Autonomous Truck Platooning. *Transportmetrica A: Transport Science* 2023; 1–37.

22 Wang C, Gan M, Zhu CA. Fault Feature Extraction of Rolling Element Bearings Based on Wavelet Packet Transform and Sparse Representation Theory. *Journal of Intelligent Manufacturing* 2018; **29**: 937–951.

23 Han S, Oh S, Jeong J. Bearing Fault Diagnosis Based on Multiscale Convolutional Neural Network Using Data Augmentation. *Journal of Sensors* 2021; **2021**(**1**): 6699637.

24 Zhang XP, Hu NQ, Hu L, *et al.* A Bearing Fault Diagnosis Method Based on Sparse Decomposition Theory. *Journal of Central South University* 2016; **23**(**8**): 1961–1969.

25 Sharma A, Amarnath M, Kankar PK. Nonlinear Dynamic Analysis of Defective Rolling Element Bearing Using Higuchi's Fractal Dimension. *Sādhanā* 2019; **44**(**4**): 76.

26 Liu N, Gu H, Wei Y, *et al.* Performance Enhancement of AlGaN-Based Deep Ultraviolet Light-Emitting Diodes by Using Stepped and Super-Lattice n-Type Confinement Layer. *Superlattices and Microstructures* 2020; **141**: 106492.

27 Feng Z, Xiong S, Cao D, *et al.* Hrs: A Hybrid Framework for Malware Detection. In Proceedings of the 2015 ACM International Workshop on International Workshop on Security and Privacy Analytics, San Antonio, TX, USA, 4 March 2015; pp. 19–26.

28 Pan S. A Survey on Transfer Learning. *IEEE Transactions on Knowledge and Data Engineering* 2010; **22**(**10**): 1345–1359.

29 Gopalan R, Li R, Chellappa R. Domain Adaptation for Object Recognition: An Unsupervised Approach. In Proceedings of the 2011 International Conference on Computer Vision, Barcelona, Spain, 6–13 November 2011; pp. 999–1006.

30 Jhuo IH, Liu D, Lee DT, *et al.* Robust Visual Domain Adaptation with Low-Rank Reconstruction. In Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition, Providence, RI, USA, 16–21 June 2012; pp. 2168–2175.

31 Pan SJ, Tsang IW, Kwok JT, *et al.* Domain Adaptation via Transfer Component Analysis. *IEEE Transactions on Neural Networks* 2010; **22**(**2**): 199–210.

32 Gretton A, Borgwardt K, Rasch M, *et al.* A Kernel Method for the Two-Sample-Problem. *Advances in Neural Information Processing Systems* 2006; **19**.

33 Sun B, Saenko K. Deep Coral: Correlation Alignment for Deep Domain Adaptation. In *Computer Vision—ECCV 2016 Workshops*; Springer International Publishing: Cham, Switzerland, 2016; pp. 443–450.

34 Song YY, Ying LU. Decision Tree Methods: Applications for Classification and Prediction. *Shanghai Archives of Psychiatry* 2015; **27**(**2**): 130.

35 Suthaharan S, Suthaharan S. Decision Tree Learning. In *Machine Learning Models and Algorithms for Big*

Data Classification: Thinking with Examples for Effective Learning; Springer: Boston, MA, USA, 2016; pp. 237–269.

36 Biau G, Scornet E. A Random Forest Guided Tour. *Test* 2016; **25**: 197–227.

37 Rigatti SJ. Random Forest. *Journal of Insurance Medicine* 2017; **47**(**1**): 31–39.

38 Peterson LE. K-Nearest Neighbor. *Scholarpedia* 2009; **4**(**2**): 1883.

39 Dai W. Design of Traffic Improvement Plan for Line 1 Baijiahu Station of Nanjing Metro. *Innovations in Applied Engineering and Technology* 2023; **10**.

40 Zhao Z, Ren P, Tang M. Analyzing the Impact of Anti-Globalization on the Evolution of Higher Education Internationalization in China. *Journal of Linguistics and Education Research* 2022; **5**(**2**): 15–31.

41 Xiong S, Yu L, Shen H, *et al.* Efficient Algorithms for Sensor Deployment and Routing in Sensor Networks for Network-Structured Environment Monitoring. In Proceedings of the 2012 IEEE INFOCOM, Orlando, FL, USA, 25–30 March 2012; pp. 1008–1016.

42 Feng Z, Cao D., Xiong S, *et al.* Method and Apparatus for File Identification. U.S. Patent No. 10284577, 7 May 2019.

43 Xiong S, Li J, Li M, *et al.* Multiple Task Scheduling for Low-Duty-Cycled Wireless Sensor Networks. In Proceedings of the 2011 IEEE INFOCOM, Shanghai, China, 10–15 April 2011; pp. 1323–1331.

44 Yu L, Li J, Cheng S, *et al.* Secure Continuous Aggregation via Sampling-Based Verification in Wireless Sensor Networks. In Proceedings of the 2011 IEEE INFOCOM, Shanghai, China, 10–15 April 2011; pp. 1763–1771.

45 Li J, Xiong S. Efficient Pr-Skyline Query Processing and Optimization in Wireless Sensor Networks. *Wireless Sensor Network* 2010; **2**(**11**): 838.

46 Wang Z, Shafieezadeh A. REAK: Reliability Analysis through Error Rate-Based Adaptive Kriging. *Reliability Engineering & System Safety* 2019; **182**: 33–45. https://doi.org/10.1016/j.ress.2018.10.004.

47 Rahimi M, Wang Z, Shafieezadeh A, *et al.* An Adaptive Kriging-Based Approach with Weakly Stationary Random Fields for Soil Slope Reliability Analysis. In Proceedings of the 2019 Eighth International Conference on Case Histories in Geotechnical Engineering, Philadelphia, PA, USA, 24–27 March 2019; pp. 148–157.

48 Wang Z, Shafieezadeh A. ESC: An Efficient Error-Based Stopping Criterion for Kriging-Based Reliability Analysis Methods. *Structural and Multidisciplinary Optimization* 2019; **59**: 1621–1637. https://doi.org/10.1007/s00158-018-2150-9.

49 Wang Z, Shafieezadeh A. Reliability-Based Bayesian Updating Using Machine Learning. In Proceedings of the 13th International Conference on Applications of Statistics and Probability in Civil Engineering (ICASP13), Seoul, Korea, 26–30 May 2019.

50 Wang Z, Shafieezadeh A. A Parallel Learning Strategy for Adaptive Kriging-Based Reliability Analysis Methods. In Proceedings of the 13th International Conference on Applications of Statistics and Probability in Civil Engineering (ICASP13), Seoul, Korea, 26–30 May 2019.

51 Hur J, Wang Z, Shafieezadeh A, *et al.* Seismic Reliability Analysis of NPP's Nonstructural Components Using Surrogate Models. Available online: https://s-space.snu.ac.kr/handle/10371/153449(accessed on 7 October 2024).

52 Darestani YM, Wang Z, Shafieezadeh A. Wind Reliability of Transmission Line Models Using Kriging-Based Methods. In Proceedings of the 13th International Conference on Applications of Statistics and Probability in Civil Engineering, ICASP13, Seoul, Korea, 26–30 May 2019.

53 Wang Z, Shafieezadeh A. Real-Time High-Fidelity Reliability Updating with Equality Information Using Adaptive Kriging. *Reliability Engineering & System Safety* 2020; **195**: 106735. https://doi.org/10.1016/j.ress.2019.106735.

54 Rahimi M, Wang Z, Shafieezadeh A, *et al.* Exploring Passive and Active Metamodeling-Based Reliability Analysis Methods for Soil Slopes: A New Approach to Active Training. *International Journal of Geomechanics* 2020; **20**: 04020009.

55 Zhang C, Wang Z, Shafieezadeh A. Value of Information Analysis via Active Learning and Knowledge

Sharing in Error-Controlled Adaptive Kriging. *IEEE Access* 2020; **8**: 51021–51034. https://doi.org/10.1109/ACCESS.2020.2980228.

56 Wang Z, Shafieezadeh A. On Confidence Intervals for Failure Probability Estimates in Kriging-Based Reliability Analysis. *Reliability Engineering & System Safety* 2020; **196**: 106758. https://doi.org/10.1016/j.ress.2019.106758.

57 Wang Z, Shafieezadeh A. Highly Efficient Bayesian Updating Using Metamodels: An Adaptive Kriging-Based Approach. *Structural Safety* 2020; **84**: 101915. https://doi.org/10.1016/j.strusafe.2019.101915.

58 Zhang C, Wang Z, Shafieezadeh A. Error Quantification and Control for Adaptive Kriging-Based Reliability Updating with Equality Information. *Reliability Engineering & System Safety* 2021; **207**: 107323. https://doi.org/10.1016/j.ress.2020.107323.

59 Wang Z, Shafieezadeh A. Metamodel-Based Subset Simulation Adaptable to Target Computational Capacities: The Case for High-Dimensional and Rare Event Reliability Analysis. *Structural and Multidisciplinary Optimization* 2021; **64**: 649–675.

60 Song C, Wang Z, Shafieezadeh A, *et al.* BUAK-AIS: Efficient Bayesian Updating with Active Learning Kriging-Based Adaptive Importance Sampling. *Computer Methods in Applied Mechanics and Engineering* 2022; **391**: 114578. https://doi.org/10.1016/j.cma.2022.114578.

61 Zhao Y, Hu H, Song C, *et al.* Predicting Compressive Strength of Manufactured-Sand Concrete Using Conventional and Metaheuristic-Tuned Artificial Neural Network. *Measurement* 2022; **194**: 110993. https://doi.org/10.1016/j.measurement.2022.110993.

62 Zhao Y, Wang Z. Subset Simulation with Adaptable Intermediate Failure Probability for Robust Reliability Analysis: An Unsupervised Learning-Based Approach. *Structural and Multidisciplinary Optimization* 2022; **65**(**6**): 172. https://doi.org/10.1007/s00158-022-03260-7.

63 Xiao X, Li Q, Wang Z. A Novel Adaptive Importance Sampling Algorithm for Bayesian Model Updating. *Structural Safety* 2022; **97**: 102230. https://doi.org/10.1016/j.strusafe.2022.102230.

64 Wang Z, Shafieezadeh A, Xiao X, *et al.* Optimal Monitoring Location for Tracking Evolving Risks to Infrastructure Systems: Theory and Application to Tunneling Excavation Risk. *Reliability Engineering & System Safety* 2022; **228**: 108781. https://doi.org/10.1016/j.ress.2022.108781.

65 Wang Z, Shafieezadeh A. Bayesian Updating with Adaptive, Uncertainty-Informed Subset Simulations: High-Fidelity Updating with Multiple Observations. *Reliability Engineering & System Safety* 2023; **230**: 108901. https://doi.org/10.1016/j.ress.2022.108901.

66 Cao M, Li Q, Wang Z. Reliability Updating with Equality Information Using Adaptive Kriging-Based Importance Sampling. *Structural and Multidisciplinary Optimization* 2023; **66**: 76. https://doi.org/10.1007/s00158-023-03492-1.

67 Ye X, Wu P, Liu A, *et al.* A Deep Learning-Based Method for Automatic Abnormal Data Detection: Case Study for Bridge Structural Health Monitoring. *International Journal of Structural Stability and Dynamics* 2023; **23**(**11**): 2350131. https://doi.org/10.1142/S0219455423501316.