Optimizations in Applied Machine Learning
Research Article | Volume 5 | Issue 5 | Feb 2025
Received: 20 Dec 2024 | Revised: 28 Jan 2025
Accepted: 11 Feb 2025 | Published Online: 13 Feb 2025

MDD-based Domain Adaptation Algorithm for Improving the Applicability of the Artificial Neural Network in Vehicle Insurance Claim Fraud Detection

Alan Wilson [1], Jiahuai Ma [2*]
[1] Intact Financial Corporation; Toronto, Ontario M5H 2N2; Canada
[2] Corresponding, University of Florida; Gainesville; Florida 32611; USA; Email: maj2@ufl.edu

Abstract: Insurance fraud detection is a critical task for insurance companies, as fraudulent claims result in financial losses and increased premiums for honest policyholders. Traditional fraud detection methods rely on rule-based approaches and manual investigation, which are limited in their ability to adapt to evolving fraud patterns. In this study, we propose a novel approach using an artificial neural network (ANN) combined with Margin Disparity Discrepancy (MDD)-based domain adaptation to improve the generalization ability of fraud detection models across different datasets. We first preprocess the data by applying K-Means clustering to segment source and target domains based on distribution differences. We then compare multiple machine learning models, including decision trees, random forests, k-nearest neighbors, and gradient-boosted decision trees, finding that ANN achieves the best performance. To further enhance generalizability, we introduce MDD-based domain adaptation, aligning feature distributions between the source and target domains. Experimental results demonstrate that the adapted ANN significantly improves fraud detection accuracy, achieving a higher F1-score and recall while reducing the false negative rate. These findings highlight the effectiveness of domain adaptation in addressing distributional shifts in fraud detection, making the proposed model a promising solution for real-world insurance fraud detection systems.

Keywords: *Vehicle insurance claim fraud detection; Machine learning; Neural network; Domain adaptation.*

1. **Introduction**

Insurance is a fundamental component of modern financial systems, providing individuals and businesses with financial protection against unforeseen risks [1][2]. Among various types of insurance, vehicle insurance plays a critical role in mitigating financial losses associated with accidents, theft, and other damages. Policyholders pay premiums to insurance companies, which, in turn, cover financial liabilities arising from claims. However, the insurance industry faces a persistent challenge in the form of fraudulent claims, which result in significant financial losses for insurers and higher premiums for honest customers. Vehicle insurance claim fraud refers to deceptive practices where individuals or groups intentionally manipulate insurance claims for financial gain. Fraudulent activities can take several forms, including exaggerated damage claims, staged accidents, and false theft reports. This fraudulent behavior creates substantial financial burdens for insurance providers, leading to increased operational costs and inflated premiums for policyholders. According to industry reports, insurance fraud costs billions of dollars annually, making fraud detection an essential component of efficient insurance operations [3][4].

Detecting fraudulent claims effectively is crucial for maintaining the financial stability of insurance companies and ensuring fairness in premium pricing [5][6]. Traditional fraud detection methods primarily rely on rule-based systems and manual investigations [7][8][9] conducted by claim adjusters. These approaches often involve predefined heuristics and business rules to flag suspicious claims. While these traditional methods have been moderately effective, they suffer from significant limitations, such as their reliance on expert knowledge [10][11][12], inability to adapt to evolving fraud patterns, and high operational costs. Consequently, there is a growing need for automated and data-driven approaches to enhance fraud detection accuracy [13][14].

With advancements in artificial intelligence (AI) and machine learning (ML) [15][16][17][18], insurers have started leveraging AI-driven techniques to improve fraud detection efficiency. AI-based fraud detection models can analyze vast amounts of claim data, identify complex patterns, and detect anomalies that may indicate fraudulent behavior. Machine learning algorithms, such as decision trees, random forests, and neural networks, have demonstrated superior performance in fraud detection by learning from historical claim data and making informed predictions [19][20][21]. Recent studies have explored various machine learning techniques for fraud detection [22][23][24], showing promising results. Decision tree-based models provide interpretable results, while deep learning methods, such as artificial neural networks (ANNs) [25][26][27], offer high accuracy by capturing intricate relationships within the data. However, one major drawback of existing machine learning models is their limited generalizability across different datasets. A model trained on one dataset (source domain) may not perform well when applied to another dataset (target domain) due to differences in data distribution. This challenge necessitates the adoption of domain adaptation techniques to enhance model adaptability.

Domain adaptation [28][29][30] is a subfield of transfer learning that aims to improve the generalization ability of machine learning models by aligning the feature distributions of the source and target domains. In the context of vehicle insurance fraud detection, domain adaptation can be used to bridge the gap between different claim datasets, allowing fraud detection models trained on one dataset to generalize effectively to another.

In this study, we propose a novel approach shown in Figure 1 leveraging Margin Disparity Discrepancy (MDD)-based domain adaptation to enhance the applicability of ANNs in vehicle insurance claim fraud detection. Our framework addresses the issue of domain shift by aligning feature distributions between the source and target domains, thereby improving model performance across different datasets. Our approach first involves data preprocessing and clustering using the K-Means algorithm. This step helps in identifying variations in claim data and allows for the formation of distinct source domain and target domain datasets with different distributions. Next, we train multiple machine learning models, including Decision Trees, Random Forest, K-Nearest Neighbors (KNN), Gradient Boosting Decision Trees (GBDT), and ANNs, on the source domain. These models are then evaluated based on their performance in predicting fraudulent claims in the target domain. Our findings reveal that among all tested models, ANN outperforms others in terms of fraud detection accuracy, making it the most suitable choice for further enhancement. To further improve the generalization of ANN model, we incorporate MDD-based domain adaptation. This involves implementing a domain adaptation mechanism that consists of two neural networks: Net-1, trained on the source domain, and Net-2, adapted for the target domain. These networks share model weights, ensuring that learned features are consistent across domains. An MDD loss function is used to minimize the difference between feature distributions extracted by neural networks. This alignment ensures that ANN model learns domain-invariant representations, making it more robust for fraud detection across different datasets.

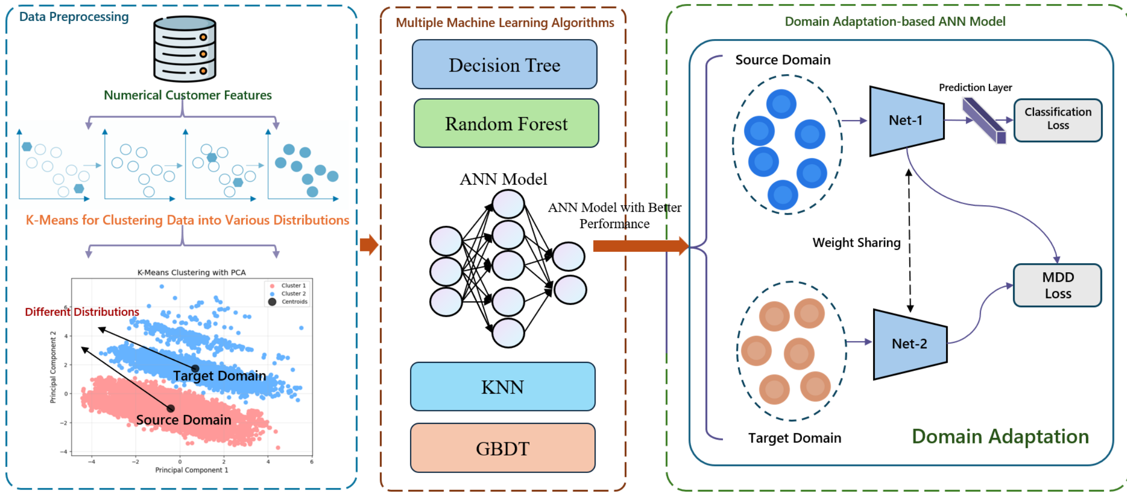

Figure 1. The process of the proposed vehicle insurance claim fraud detection method using MDD-based domain adaptation.

2. **Literature Review**

A. Insurance fraud detection

Insurance fraud detection has garnered significant attention, leading to the development of various machine learning methodologies aimed at enhancing detection accuracy and efficiency due to their excellent performance in many tasks [31][32][33][34]. For instance, Roy et al. introduced a framework that combines feature selection with ensemble learning techniques to detect fraudulent claims, achieving notable improvements in accuracy and a reduction in false positives [35]. Gangadhar et al. developed a Chaotic Variational Autoencoder-based one-class classifier tailored for insurance fraud detection, demonstrating significant improvements in identifying fraudulent transactions [36]. Additionally, Asgarian et al. introduced AutoFraudNet, a multimodal network designed to detect fraud in the auto insurance industry by integrating various data modalities to enhance detection performance [37]. Gupta et al. applied a Markov model integrated with machine learning techniques for fraud detection in health insurance, achieving high accuracy and F1-scores, thereby demonstrating the model's effectiveness [38]. However, while these studies contribute valuable insights into insurance fraud detection, they do not consider the issue of varying data distributions across different datasets, which may limit their effectiveness in real-world applications.

B. Domain adaptation

Pan et al. categorized transfer learning into three primary types based on variations in domains and tasks: inductive, transductive, and unsupervised transfer learning [39]. Inductive transfer learning occurs when the source and target tasks differ, regardless of whether they belong to the same domain. This type often utilizes labeled data from the source domain, but it also requires at least some labeled data in the target domain for training. Transductive transfer learning, on the other hand, maintains the same task across different domains but has labeled data only in the source domain. During training, a portion of the target domain's unlabeled data is used to estimate its marginal probability distribution. Lastly, unsupervised transfer learning involves differences in both tasks and domains, similar to inductive learning, but without any labeled data in either domain, relying entirely on unsupervised techniques.

From a methodological perspective, domain adaptation can be broadly divided into two categories based on their structural approaches: shallow and deep methods. Shallow domain adaptation methods, as referenced in [40][41][42], primarily employ instance-based and feature-based strategies to align domain distributions. A common approach is minimizing the distance between domains using metrics such as the Wasserstein metric, correlation alignment (CORAL), Kullback-Leibler (KL) divergence, and contrastive domain discrepancy (CDD). In contrast, deep domain adaptation methods [43][44][45] leverage neural networks, typically incorporating convolutional, autoencoder, or adversarial architectures to bridge the domain gap. These approaches often integrate distance metrics at different layers of dual-network structures, where one network processes the source domain while the other handles the target domain, enabling the measurement and reduction of discrepancies in feature representations across layers.

3. **Method**

A. Dataset descrption and preprocessing

Our study utilizes a publicly available dataset from Kaggle, which contains 33 features aimed at detecting fraudulent claims in vehicle insurance. The primary objective of this research is to develop an effective fraud detection model, where we consider the "FraudFound_P" feature as the target variable. The dataset includes various features representing different aspects of insurance claims, such as claim-related attributes, policyholder information, and accident details. Some example features include Month, WeekOfMonth, DayOfWeek, among others. The target variable "FraudFound_P" indicates whether a claim is fraudulent (1) or legitimate (0). In our dataset, fraud accounts for approximately 5.99% of the total claims, while 94.01% of the claims are legitimate. Figure 2 and Figure 3 illustrate the distribution of several numerical and categorical features, providing insights into data characteristics and variability across different claim attributes.

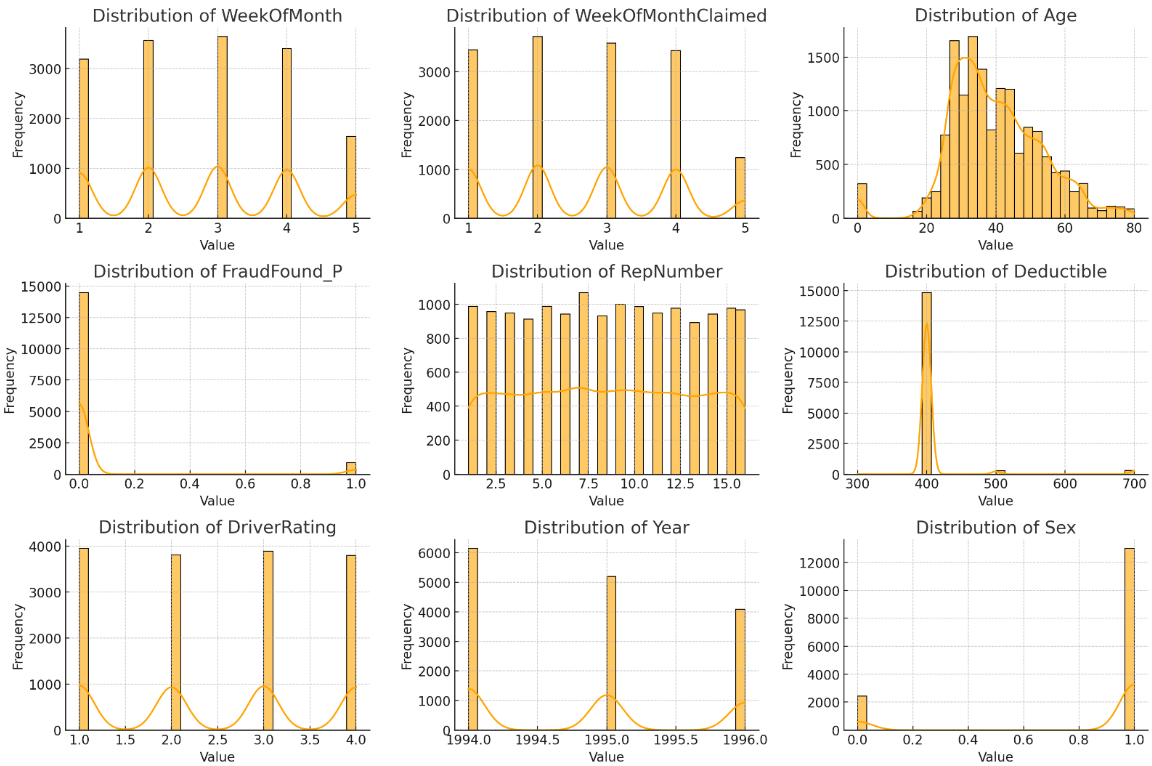

Figure 2. The distribution of some numerical features.

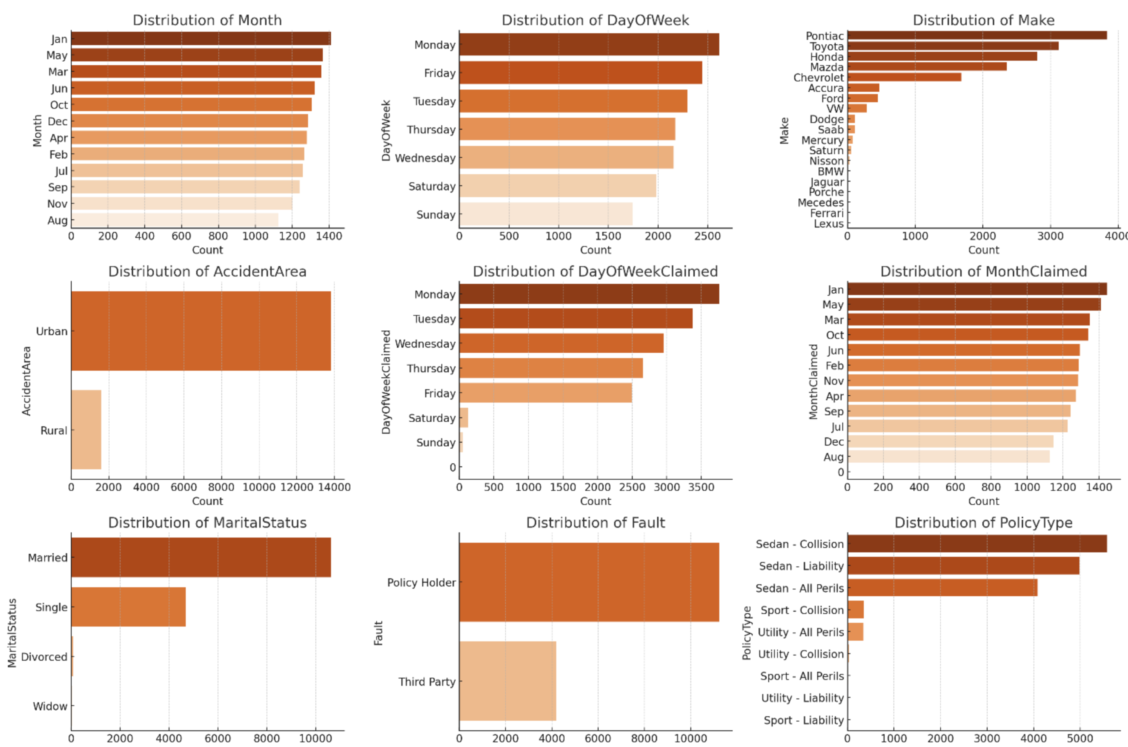

Figure 3. The distribution of some categorical features.

To ensure the dataset is suitable for the domain adaptation task, we first split the data into a source domain and a target domain. Determining the optimal number of clusters (K) was crucial in this step. We applied the Elbow Method and Silhouette Scores, as shown in Figure 4, to analyze the clustering performance across different values of K. The results indicated that K=2 provided the best clustering effectiveness, leading us to divide the data into two clusters. Figure 5 illustrates the PCA-based data distribution, where the left plot shows the original dataset projected onto two principal components, while the right plot presents the segmented clusters based on K-means. The two identified clusters will serve as our source and target domains, allowing for further domain adaptation processing. Thereinto, cluster 1 and cluster 2 are used as the source domain (training data) and target domain (validation and testing data), respectively in our study. Additionally, we recognized that the dataset contains many irrelevant features which do not contribute meaningfully to fraud detection. To address this, we applied a random forest model for feature selection, the results of which will be discussed in the later experimental section. Furthermore, considering that the dataset is highly imbalanced, with the number of fraud cases significantly lower than non-fraud cases, we performed a downsampling operation to balance the dataset before proceeding with model training.

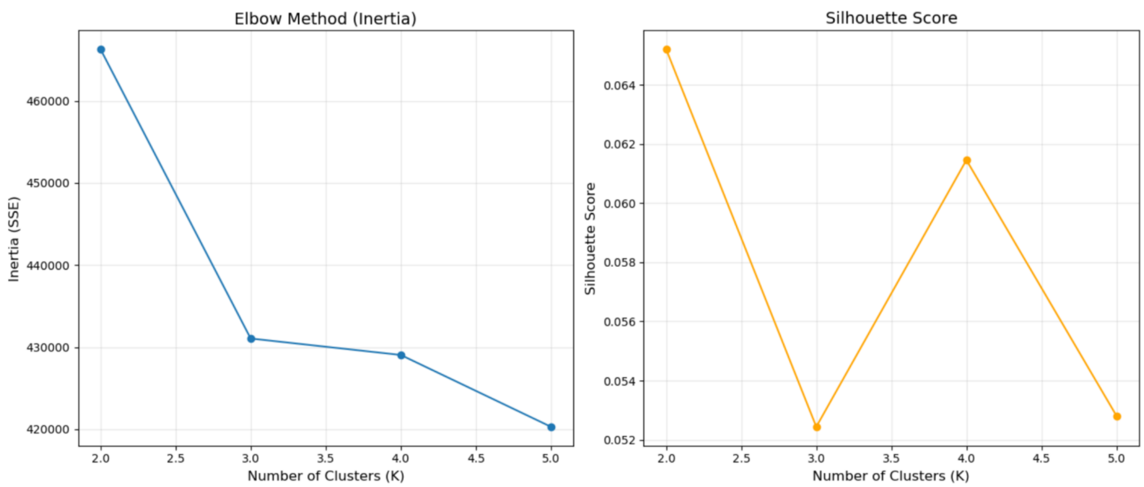

Figure 4. The curves of Elbow method and Silhouette scores used for determining k value.

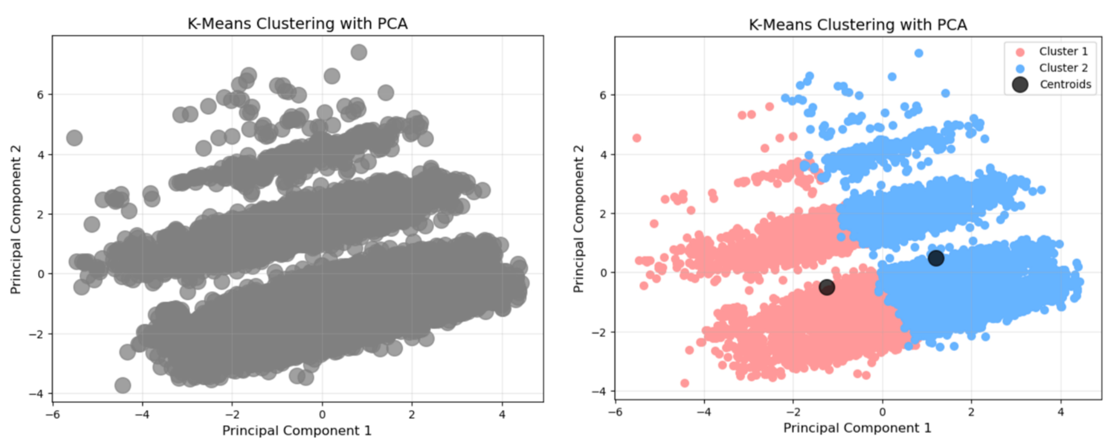

Figure 5. The PCA-based original distribution (left) and distribution with segmented clusters (right) of the data.

B. Machine learning models

To determine the optimal model for integration with the domain adaptation algorithm, we first trained five different machine learning models on the source domain and then tested their performance directly on the target domain. These models include 1) Decision tree, 2) Random Forest, 3) K-Nearest Neighbors, 4) Gradient-boosted decision trees, and 5) Artificial neural network. Below is an introduction to the basic principles of each model.

1. Decision tree

A decision tree [46] is a tree-structured supervised learning algorithm commonly used for classification and regression tasks. It recursively splits the dataset based on feature values to create a hierarchical structure of decision nodes and leaf nodes. The splits are chosen to maximize information gain using metrics such as Gini impurity or entropy for classification and mean squared error for regression. Decision trees are easy to interpret and require minimal preprocessing, but they are prone to overfitting, especially when the tree depth is not properly controlled.

2. Random forest

Random forest [47] is an ensemble learning method that builds multiple decision trees and aggregates their predictions to improve model performance and robustness. Each tree is trained on a randomly sampled subset of the data, and feature selection is performed randomly at each node to increase diversity. The final prediction is determined by majority voting for classification or averaging for regression. Random forests are more resistant to overfitting compared to single decision trees and perform well with high-dimensional data. However, they can be computationally expensive when dealing with large datasets.

3. K-Nearest neighbors

KNN [48] is a non-parametric, instance-based learning algorithm that classifies a sample by analyzing the labels of its k nearest neighbors in feature space. It relies on distance metrics such as Euclidean distance or Manhattan distance to determine similarity between points. KNN is simple to implement and works well with well-separated classes but can be computationally expensive when the dataset is large. Additionally, it is sensitive to irrelevant features and requires careful selection of the k parameter for optimal performance.

4. Gradient-boosted decision trees

GBDT [49] is an ensemble learning technique that builds multiple decision trees sequentially, where each tree corrects the errors of the previous one. The model optimizes a loss function using gradient descent, making it highly effective for structured data. Unlike random forest, which builds trees independently, GBDT learns iteratively by adjusting the weights of samples that were misclassified in previous trees. It provides high accuracy and is widely used in various machine learning applications. However, it is more prone to overfitting than random forests and requires careful tuning of hyperparameters such as learning rate and number of trees.

5. Artificial neural network

ANN [50][51]is a deep learning model inspired by the structure of biological neural networks. It consists of layers of interconnected neurons, where each neuron processes input data through weighted connections and activation functions. ANN can capture complex patterns and non-linear relationships in data, making it powerful for tasks such as fraud detection. Training an ANN involves adjusting the weights using backpropagation and optimization algorithms such as stochastic gradient descent (SGD) or Adam. While ANNs offer high flexibility and accuracy, they require large amounts of data and computational resources, and they are less interpretable compared to decision trees or random forests.

C. Domain adaptation-based artificial neural network model

Domain adaptation is a specialized area within transfer learning that focuses on enhancing model performance when there are distributional differences between the source domain (training data) and the target domain (test data). Unlike traditional machine learning methods that assume identical distributions for training and testing, domain adaptation mitigates these discrepancies by aligning feature distributions across domains. This adaptation ensures that models maintain strong predictive performance even when applied to new or varied data. Domain adaptation techniques can be broadly categorized into shallow and deep approaches. Shallow methods typically employ statistical techniques to minimize discrepancies between domains, utilizing metrics such as Kullback-Leibler (KL) divergence to achieve feature alignment. In contrast, deep domain adaptation leverages neural networks to extract transferable features, often integrating adversarial learning or dual-network architectures to enhance domain alignment.

This technique is particularly beneficial in situations where labeled data in the target domain is limited or unavailable. By leveraging labeled data from the source domain alongside unlabeled or sparsely labeled target data, domain adaptation improves generalization and reduces the need for extensive data collection and annotation. It has found widespread applications in fields such as image recognition, natural language processing, and biometric authentication, effectively addressing challenges arising from domain shifts. Figure 6 provides a conceptual illustration of the domain adaptation process.

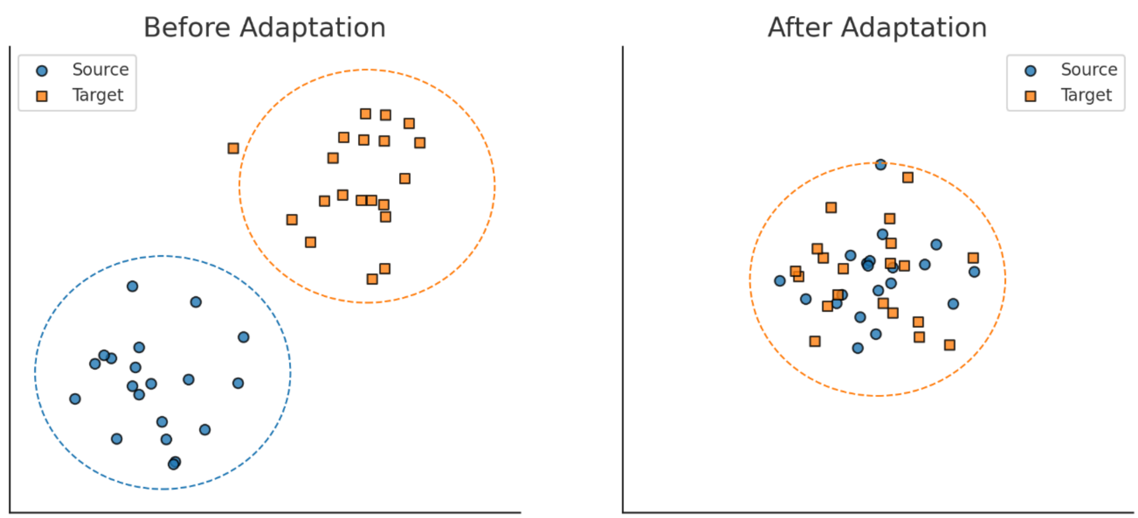

Figure 6. The schematic of the idea related to domain adaptation.

In our experiments, we observed that ANN outperformed the other models in terms of predictive accuracy and generalization ability. Consequently, we selected ANN as the base model for integrating with the domain adaptation framework. The ANN used in our study consists of five fully connected layers. The first layer has 128 neurons and serves as the input layer, followed by three hidden layers containing 64, 32, and 16 neurons, respectively. Each of these layers employs the ReLU activation function, which helps in capturing complex patterns while mitigating the vanishing gradient problem. The final layer consists of a single neuron with a sigmoid activation function, making it suitable for binary classification tasks. To optimize the network, we used an adaptive gradient-based optimizer with a learning rate of 0.01, which helps efficiently update the model weights during training. The loss function chosen for training is binary cross-entropy, a standard choice for classification problems where the output is a probability score. The model was trained with a batch size of 256 for 30 epochs, ensuring stable convergence and effective learning from the source domain data. To incorporate domain adaptation, we extracted the output features from the layer with 64 neurons, as this layer provides an optimal intermediate representation, balancing abstraction and detailed feature retention. These extracted features were then used for domain adaptation alignment between the source and target domains.

For distribution alignment, we employed the MDD loss function. MDD is a statistical metric that measures the difference between two probability distributions by comparing their mean embeddings in a reproducing kernel Hilbert space. By minimizing MDD loss, the model learns domain-invariant feature representations, reducing the distributional gap between the source and target domains. This enhances the model's ability to generalize across domains, improving its performance on the target domain despite differences in data distribution. The total loss function is shown in equation (1).

$$L_{Total} = L_{CE} + \lambda L_{MDD} \tag{1}$$

where $\mathrm{L_{CE}}$ represents the binary cross-entropy loss, which is used for classification. $\mathrm{L_{MDD}}$ denotes the MDD loss, which aligns the feature distributions of the source and target domains. In this study, the λ which corresponds to the weight of MDD loss, was automatically adjusted during training following a sigmoid-based schedule, gradually increasing from 0.0 to 1.0 over 1000 steps with a speed factor of 1.0.

4. **Results and Discussion**

A. The feature importance of the random forest

As we previously considered using a random forest model for feature selection, Figure 7 presents the importance of features produced by the random forest model. The feature importance scores indicate how much each feature contributes to the model's predictive performance, helping to identify the most relevant attributes for fraud detection. From the results, policy number emerges as the most important feature, which is likely due to specific patterns or anomalies within policy numbers that may be correlated with fraudulent claims. However, relying on such an identifier-based feature could lead to data leakage, as policy numbers should not inherently determine fraud. Age and rep number also show high importance, which seems reasonable, as older policyholders may exhibit different claim behaviors compared to younger ones, and rep number may be associated with claim history or adjustments by insurance representatives. Temporal features such as month claimed, month, day of week, and week of month also rank highly in importance, suggesting that fraud occurrence may have seasonal or time-dependent patterns. This aligns with previous studies indicating that fraud claims often peak during certain periods, possibly due to strategic filing behaviors. Other notable features include fault, driver rating, past number of claims, and number of supplements, all of which relate to driving behavior, past claims, and additional policy elements. These features align well with domain knowledge, as fraud is often linked to claim history and driver-related characteristics. On the other hand, some features with relatively low importance, such as witness present, agent type, and vehicle category, may not have strong predictive power in fraud detection. This suggests that these attributes either contain redundant information or lack a significant relationship with fraudulent claims.

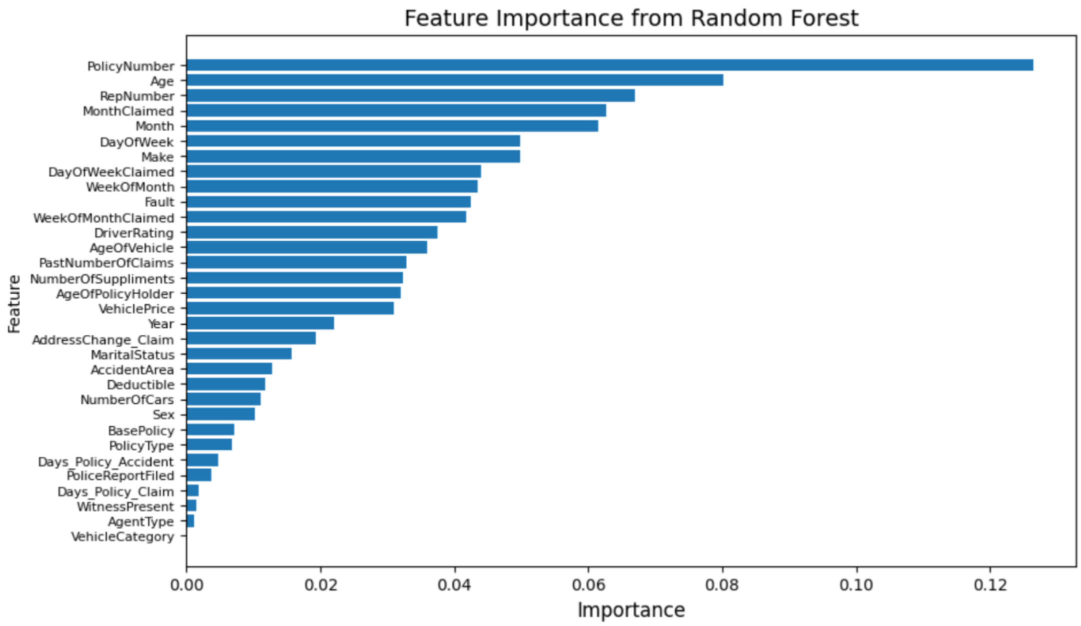

Figure 7. The importance of features produced from the random forest model.

B. The performance of multiple machine learning models for direct prediction in target domain

To determine the most effective model for fraud detection before integrating domain adaptation, we evaluated multiple machine learning models, including decision trees, random forests, KNN, GBDT, and ANN. Each model was trained on the source domain and directly tested on the target domain to assess its generalization ability. Table 1 presents the performance of these models based on accuracy, F1-score, precision, and recall. The results indicate that ANN outperformed other models, particularly in terms of F1-score and precision. Decision tree and random forest models showed relatively high recall but extremely low precision, meaning they flagged a significant number of fraudulent cases but with a high false positive rate. KNN and GBDT demonstrated better overall balance, but their F1-scores were still considerably lower than ANN. Figure 8 illustrates the training accuracy and loss curves of ANN model. The training and validation accuracy gradually improve over the epochs, showing that the model successfully learns from the data. Meanwhile, the loss curve indicates a rapid convergence within the first few epochs, stabilizing at a low value, suggesting effective optimization.

To provide a clearer comparison, Figure 9 visualizes the performance metrics for each model. ANN achieves the highest F1-score and precision, which are crucial in fraud detection since they indicate the model's ability to correctly identify fraudulent claims while minimizing false positives. While decision tree and random forest models exhibit higher recall, their low precision makes them less reliable for real-world deployment, as they would result in a large number of false alarms. Figure 10 presents the ROC curves for different models, where the area under the curve (AUC) serves as an indicator of the model's discrimination ability. ANN achieves the highest AUC of 0.55, outperforming the other models, which remain around 0.50 to 0.53, indicating that they struggle to differentiate fraudulent claims from legitimate ones effectively. Finally, Figure 11 displays the confusion matrices for all models, highlighting their classification behavior. The ANN model correctly identifies a significant number of fraud cases compared to the others, with a better balance between false positives and false negatives. In contrast, traditional machine learning models, especially decision trees and random forests, classify most instances as non-fraudulent, failing to capture the complex patterns of fraud detection.

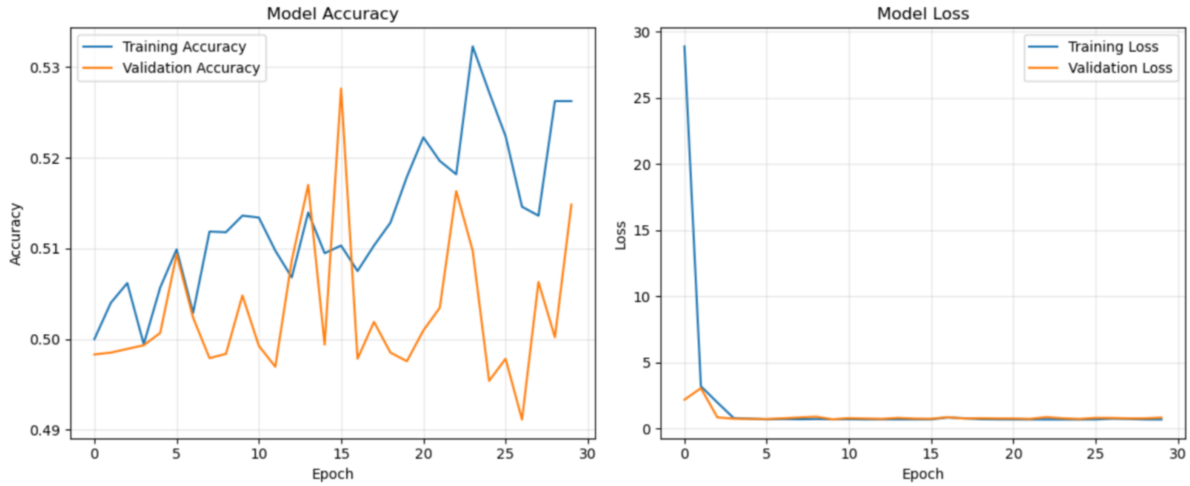

Figure 8. The training accuracy (left) and loss (right) of ANN model.

Table 1. The performance of models evaluated by different metrics.

Metrics	Accuracy	F1-score	Recall	Preicision
Decision tree	0.5018	0.0081	0.0041	0.8824
Random forest	0.5035	0.0145	0.0073	0.9643
K-Nearest neighbors	0.5141	0.1485	0.0847	0.5998
Gradient-boosted decision tree	0.5270	0.1627	0.0919	0.7077
Artificial neural network	0.5328	0.4781	0.4280	0.5415

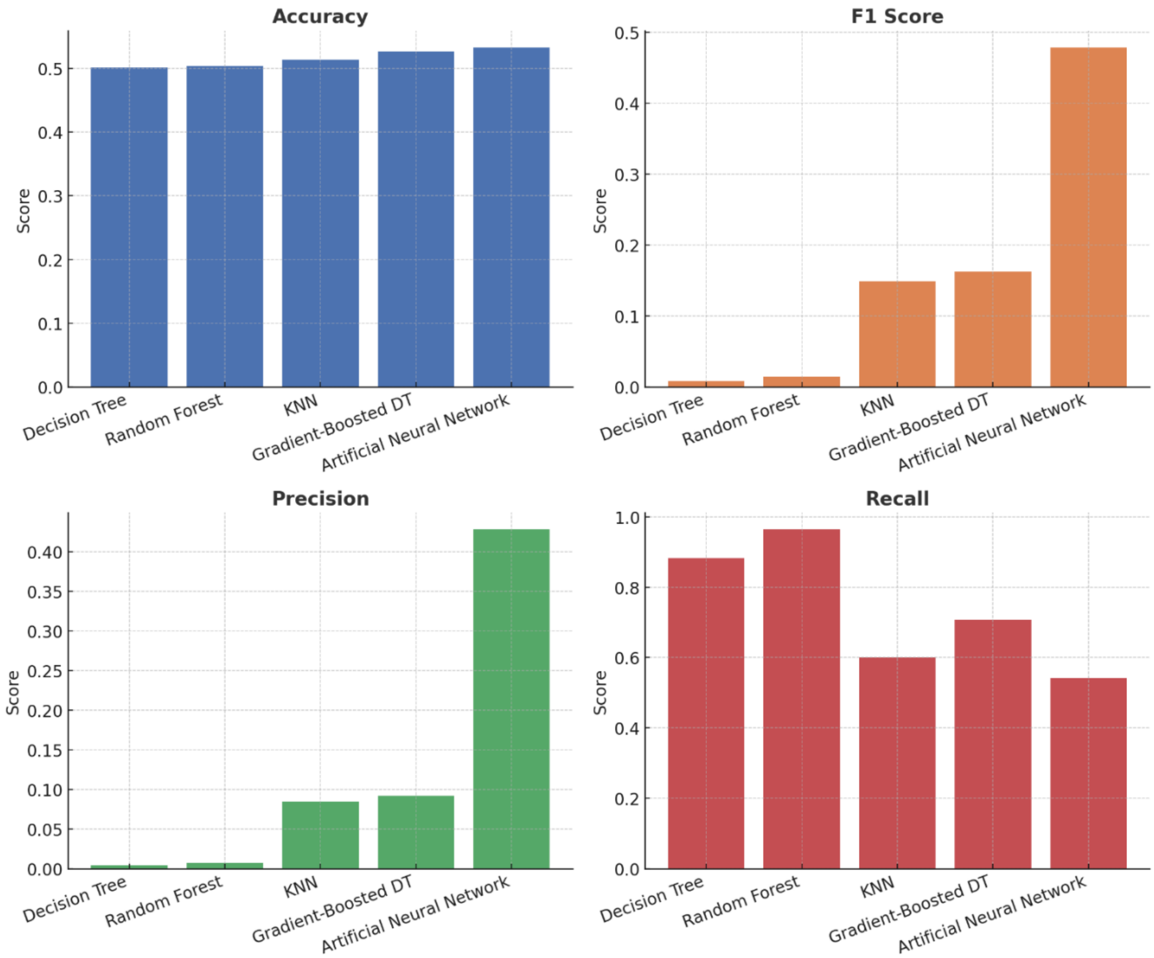

Figure 9. The visualization of performance comparison based on various models.

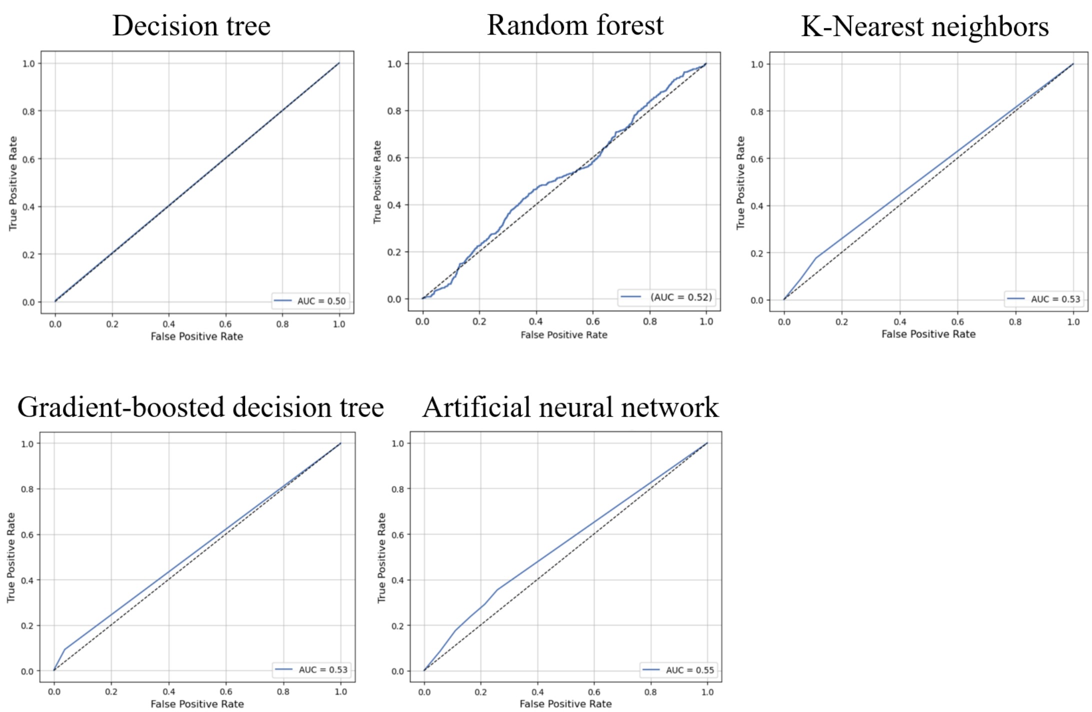

Figure 10. The ROC curves of different models.

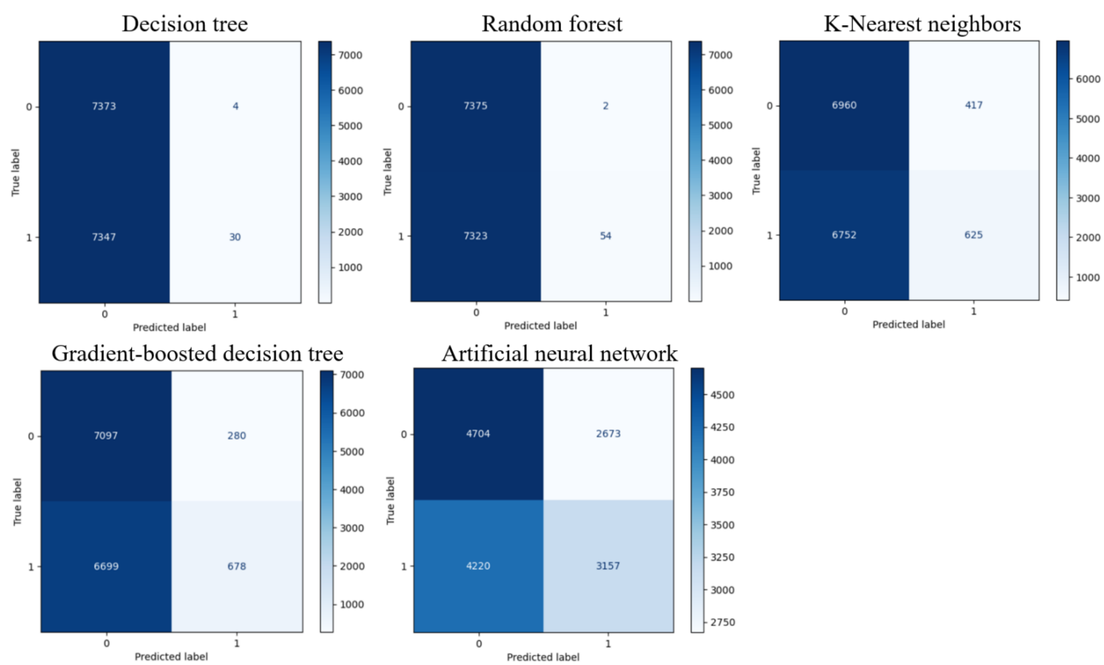

Figure 11. Confusion matrices of different models.

C. The performance of domain adaptation-based artificial neural network in target domain

To evaluate the effectiveness of domain adaptation in improving fraud detection, we compared the performance of the original ANN model with the domain adaptation-based ANN. The results demonstrate significant improvements in key evaluation metrics after applying domain adaptation.

Figure 12 presents the discriminator loss during training and the feature distribution after applying domain adaptation. The left plot shows the discriminator loss fluctuating at the beginning of training before stabilizing, indicating the model's learning process in aligning the source and target distributions. The right plot visualizes the transformed feature space after applying dimensionality reduction, showing a more overlapped distribution of the source and target domain data, suggesting that the adaptation process successfully reduced the domain shift.

Table 2 provides a quantitative comparison of model performance. The domain adaptation-based ANN achieved an accuracy of 0.7351, compared to 0.5328 for the original ANN. The F1-score also saw a significant increase from 0.4781 to 0.6499, indicating improved balance between precision and recall. The precision improved slightly from 0.4280 to 0.4918, while the recall increased dramatically from 0.5415 to 0.9578. This suggests that domain adaptation enabled the model to correctly identify a significantly larger portion of fraudulent claims while maintaining a reasonable false positive rate. Figure 13 further illustrates these improvements using confusion matrices. The original ANN model misclassified a large number of fraudulent claims, with many fraud cases being classified as non-fraudulent. In contrast, the domain adaptation-based ANN significantly reduced false negatives, correctly identifying a greater proportion of fraud cases. The number of false positives also decreased, highlighting the model's improved reliability in distinguishing fraudulent and legitimate claims.

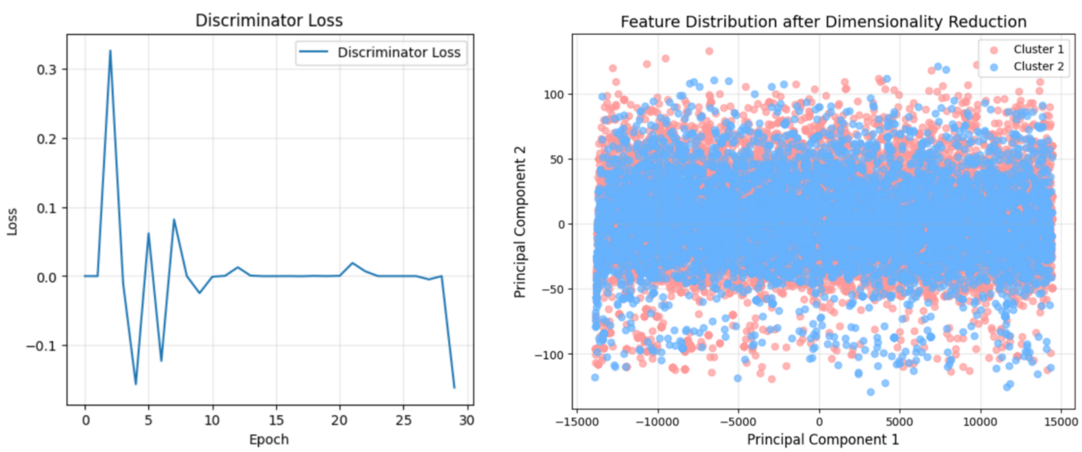

Figure 12. The discriminator loss during ANN training process (left) and transformed distributions of source and target domains (right).

Table 2. The performance of ANN model with and without domain adaptation.

Metrics	Accuracy	F1-score	Recall	Preicision
Original artificial neural network	0.5328	0.4781	0.4280	0.5415
Domain adaptation-based artificial neural network	0.7351	0.6499	0.4918	0.9578

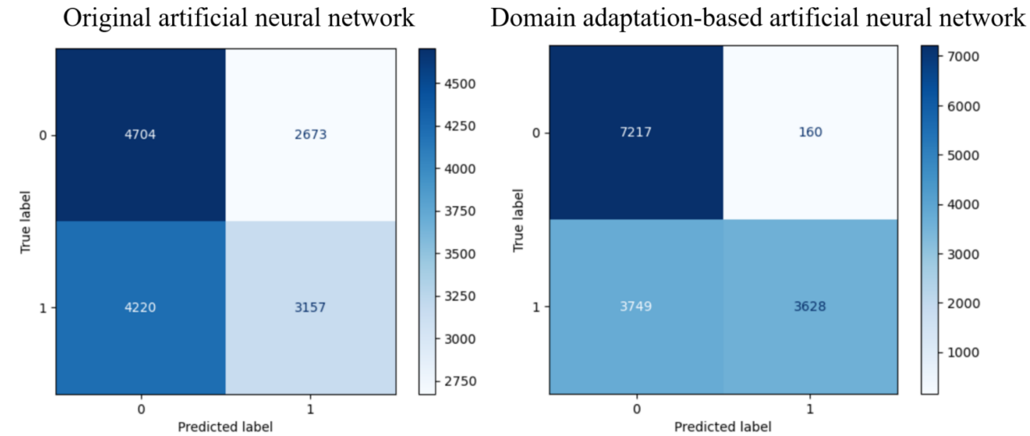

Figure 13. The confusion matrices of original and domain adaptation-based artificial neural networks.

D. Discussion

While the results demonstrate significant improvements in fraud detection using domain adaptation, several limitations should be acknowledged. One major limitation is the reliance on labeled data from the source domain while assuming no or very limited labeled data in the target domain. Although domain adaptation techniques attempt to bridge the distributional gap, the model's performance still heavily depends on the quality and representativeness of the source domain data. If the source domain data is biased or lacks diversity, the model may struggle to generalize effectively, even after adaptation. Another challenge is the choice of hyperparameters in both ANN and the domain adaptation framework. The adaptation strength, learning rate, and the balancing coefficient for MDD loss all play critical roles in model performance. While we employed an adaptive balancing parameter for MDD loss, further fine-tuning may be necessary to optimize the trade-off between classification accuracy and domain alignment. Additionally, the ANN architecture itself, including the number of layers and neurons, could be further refined to better capture complex fraud patterns.

The evaluation also reveals that, despite improvements in recall, precision remains relatively low. This indicates that the model still produces a considerable number of false positives, which could lead to inefficiencies in real-world fraud detection systems. While reducing false negatives is critical to identifying fraudulent claims, a high false positive rate could increase operational costs for insurance companies, requiring further manual verification. Future work could explore alternative domain adaptation strategies, such as adversarial training or contrastive learning, to better distinguish between fraudulent and legitimate claims.

5. **Conclusion**

This study investigates the application of domain adaptation in fraud detection, addressing the challenge of distributional differences between datasets. Our experiments demonstrate that conventional machine learning models struggle to generalize when trained on one dataset and tested on another. By integrating an ANN with MDD-based domain adaptation, we improve the model's ability to detect fraudulent claims across different domains. The proposed approach effectively reduces the domain shift by aligning feature distributions, resulting in significant performance gains in recall and overall detection accuracy. Despite these improvements, the model still faces challenges related to false positives, which could increase operational costs in real-world applications. Future research should explore advanced domain adaptation techniques, such as adversarial learning, and incorporate dynamic adaptation strategies to enhance fraud detection robustness. Overall, this study demonstrates that domain adaptation is a viable solution for improving fraud detection models, providing insurance companies with a more reliable and adaptable fraud detection framework.

Funding
Not applicable

Author Contributions
Alan Wilson contributed to the theoretical analysis, algorithm development, and implementation. Jiahuai Ma supervised the research, refined the methodology, and contributed to data analysis and manuscript writing. Both authors approved the final manuscript.

Institutional Reviewer Board Statement
Not applicable

Informed Consent Statement
Not applicable

Data Availability Statement
The data supporting the findings of this study are available from the corresponding author upon request.

Conflict of Interest
The authors declare no conflict of interest.

References

[1] Ewold F. Insurance and risk. The Foucault effect: Studies in governmentality. 1991;197210:201-2.

[2] Spence M, Zeckhauser R. Insurance, information, and individual action. In Uncertainty in economics 1978 Jan 1 (pp. 333-343). Academic Press.

[3] Derrig RA. Insurance fraud. Journal of Risk and Insurance. 2002 Sep;69(3):271-87.

[4] Viaene S, Dedene G. Insurance fraud: Issues and challenges. The Geneva Papers on Risk and Insurance-Issues and Practice. 2004 Apr 1;29:313-33.

[5] Sithic HL, Balasubramanian T. Survey of insurance fraud detection using data mining techniques. arXiv preprint arXiv:1309.0806. 2013 Sep 3.

[6] Abdallah A, Maarof MA, Zainal A. Fraud detection system: A survey. Journal of Network and Computer Applications. 2016 Jun 1;68:90-113.

[7] Zhou Z, Wu J, Cao Z, She Z, Ma J, Zu X. On-Demand Trajectory Prediction Based on Adaptive Interaction Car Following Model with Decreasing Tolerance. In2021 International Conference on Computers and Automation (CompAuto) 2021 Sep 7 (pp. 67-72). IEEE.

[8] Ding X, Wing JJ, Gibbs BH, Drum E, Gutierrez D, Albala B, Boden-Albala B. Abstract WP183: Social Activities Are Positively Associated With Medication Adherence In Stroke Survivors. Stroke. 2023 Feb;54(Suppl_1):AWP183-.

[9] Dai W. Evaluation and improvement of carrying capacity of a traffic system. Innovations in Applied Engineering and Technology. 2022 Nov 22:1-9.

[10] Dai W. Safety evaluation of traffic system with historical data based on Markov process and deep-reinforcement learning. Journal of Computational Methods in Engineering Applications. 2021 Oct 21:1-4.

[11] 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.

[12] Lei J. Efficient Strategies on Supply Chain Network Optimization for Industrial Carbon Emission Reduction. arXiv preprint arXiv:2404.16863. 2024 Apr 17.

[13] Badr MM, Ibrahem MI, Kholidy HA, Fouda MM, Ismail M. Review of the data-driven methods for electricity fraud detection in smart metering systems. Energies. 2023 Mar 19;16(6):2852.

[14] Tran PH, Tran KP, Huong TT, Heuchenne C, HienTran P, Le TM. Real time data-driven approaches for credit card fraud detection. InProceedings of the 2018 international conference on e-business and applications 2018 Feb 23 (pp. 6-9).

[15] Zhu Y, Zhao Y, Song C, Wang Z. Evolving reliability assessment of systems using active learning-based surrogate modelling. Physica D: Nonlinear Phenomena. 2024 Jan 1;457:133957.

[16] Ye X, Luo K, Wang H, Zhao Y, Zhang J, Liu A. An advanced AI-based lightweight two-stage underwater structural damage detection model. Advanced Engineering Informatics. 2024 Oct 1;62:102553.

[17] Chen X, Wang M, Zhang H. Machine Learning-based Fault Prediction and Diagnosis of Brushless Motors. Engineering Advances. 2024 Aug 6;4(3).

[18] Chen X, Gan Y, Xiong S. Optimization of Mobile Robot Delivery System Based on Deep Learning. Journal of Computer Science Research. 2024;6(4):51-65.

[19] Zhang H, Zhu D, Gan Y, Xiong S. End-to-End Learning-Based Study on the Mamba-ECANet Model for Data Security Intrusion Detection. Journal of Information, Technology and Policy. 2024 Sep 7:1-7.

[20] Gan Y, Ma J, Xu K. Enhanced E-Commerce Sales Forecasting Using EEMD-Integrated LSTM Deep Learning Model. Journal of Computational Methods in Engineering Applications. 2023 Nov 11:1-1.

[21] Zhang G, Zhou T. Finite Element Model Calibration with Surrogate Model-Based Bayesian Updating: A Case Study of Motor FEM Model. Innovations in Applied Engineering and Technology. 2024 Sep 30:1-3.

[22] Motie S, Raahemi B. Financial fraud detection using graph neural networks: A systematic review. Expert Systems with Applications. 2024 Apr 15;240:122156.

[23] Chatterjee P, Das D, Rawat DB. Digital twin for credit card fraud detection: Opportunities, challenges, and fraud detection advancements. Future Generation Computer Systems. 2024 Apr 30.

[24] Mienye ID, Jere N. Deep learning for credit card fraud detection: A review of algorithms, challenges, and solutions. IEEE Access. 2024 Jul 11.

[25] Lu X, Jiang Q, Shen Y, Lin X, Xu F, Zhu Q. Enhanced residual convolutional domain adaptation network with CBAM for RUL prediction of cross-machine rolling bearing. Reliability Engineering & System Safety. 2024 May 1;245:109976.

[26] Ding X, Meng Y, Xiang L, Boden-Albala B. Stroke recurrence prediction using machine learning and segmented neural network risk factor aggregation. Discover Public Health. 2024 Oct 7;21(1):119.

[27] Jihu L. Green Supply Chain Management Optimization Based on Chemical Industrial Clusters. arXiv preprint arXiv:2406.00478. 2024 Jun 1.

[28] Farahani A, Voghoei S, Rasheed K, Arabnia HR. A brief review of domain adaptation. Advances in data science and information engineering: proceedings from ICDATA 2020 and IKE 2020. 2021:877-94.

[29] Ben-David S, Blitzer J, Crammer K, Pereira F. Analysis of representations for domain adaptation. Advances in neural information processing systems. 2006;19.

[30] Wang M, Deng W. Deep visual domain adaptation: A survey. Neurocomputing. 2018 Oct 27;312:135-53.

[31] Gan Y, Zhu D. The Research on Intelligent News Advertisement Recommendation Algorithm Based on Prompt Learning in End-to-End Large Language Model Architecture. Innovations in Applied Engineering and Technology. 2024 Aug 30:1-9.

[32] Chen X, Zhang H. Performance Enhancement of AlGaN-based Deep Ultraviolet Light-emitting Diodes with AlxGa1-xN Linear Descending Layers. Innovations in Applied Engineering and Technology. 2023 Oct 31:1-0.

[33] Wang Z, Zhao Y, Song C, Wang X, Li Y. A new interpretation on structural reliability updating with adaptive batch sampling-based subset simulation. Structural and Multidisciplinary Optimization. 2024 Jan;67(1):7.

[34] Wenjun D, Fatahizadeh M, Touchaei HG, Moayedi H, Foong LK. Application of six neural network-based solutions on bearing capacity of shallow footing on double-layer soils. Steel and Composite Structures. 2023;49(2):231-44.

[35] Roy R, George KT. Detecting insurance claims fraud using machine learning techniques. In2017 international conference on circuit, power and computing technologies (ICCPCT) 2017 Apr 20 (pp. 1-6). IEEE.

[36] Gangadhar KS, Kumar BA, Vivek Y, Ravi V. Chaotic variational auto encoder based one class classifier for insurance fraud detection. arXiv preprint arXiv:2212.07802. 2022 Dec 15.

[37] Asgarian A, Saha R, Jakubovitz D, Peyre J. AutoFraudNet: A Multimodal Network to Detect Fraud in the Auto Insurance Industry. arXiv preprint arXiv:2301.07526. 2023 Jan 15.

[38] Gupta RY, Mudigonda SS, Baruah PK, Kandala PK. Markov model with machine learning integration for fraud detection in health insurance. arXiv preprint arXiv:2102.10978. 2021 Feb 11.

[39] S. Pan, A survey on transfer learning. IEEE Trans. Knowl. Data Eng. 22(10), 1345–1359 (2010)

[40] Gopalan, R., Li, R., & Chellappa, R. (2011, November). Domain adaptation for object recognition: An unsupervised approach. In 2011 international conference on computer vision (pp. 999-1006). IEEE.

[41] Jhuo, I. H., Liu, D., Lee, D. T., & Chang, S. F. (2012, June). Robust visual domain adaptation with low-rank reconstruction. In 2012 IEEE conference on computer vision and pattern recognition (pp. 2168-2175). IEEE.

[42] Pan, S. J., Tsang, I. W., Kwok, J. T., & Yang, Q. (2010). Domain adaptation via transfer component analysis. IEEE transactions on neural networks, 22(2), 199-210.

[43] Ajakan, H., Germain, P., Larochelle, H., Laviolette, F., & Marchand, M. (2014). Domain-adversarial neural networks. arXiv preprint arXiv:1412.4446.

[44] Ganin, Y., & Lempitsky, V. (2015, June). Unsupervised domain adaptation by backpropagation. In International conference on machine learning (pp. 1180-1189). PMLR.

[45] Glorot, X., Bordes, A., & Bengio, Y. (2011). Domain adaptation for large-scale sentiment classification: A deep learning approach. In Proceedings of the 28th international conference on machine learning (ICML-11) (pp. 513-520).

[46] Quinlan JR. Learning decision tree classifiers. ACM Computing Surveys (CSUR). 1996 Mar 1;28(1):71-2.

[47] Cutler A, Cutler DR, Stevens JR. Random forests. Ensemble machine learning: Methods and applications. 2012:157-75.

[48] Kramer O, Kramer O. K-nearest neighbors. Dimensionality reduction with unsupervised nearest neighbors. 2013:13-23.

[49] Zhang Z, Jung C. GBDT-MO: gradient-boosted decision trees for multiple outputs. IEEE transactions on neural networks and learning systems. 2020 Aug 4;32(7):3156-67.

[50] Zhao Y, Dai W, Wang Z, Ragab AE. Application of computer simulation to model transient vibration responses of GPLs reinforced doubly curved concrete panel under instantaneous heating. Materials Today Communications. 2024 Mar 1;38:107949.

[51] Hao Y, Chen Z, Sun X, Tong L. Planning of truck platooning for road-network capacitated vehicle routing problem. Transportation Research Part E: Logistics and Transportation Review. 2025 Feb 1;194:103898.