Journal of Information Systems and Informatics
Vol. 7, No. 1, March 2025 e-ISSN: 2656-4882 p-ISSN: 2656-5935
DOI: 10.51519/journalisi.v7i1.958 Published By DRPM-UBD

Hybrid Unsupervised Machine Learning for Insurance Fraud Detection: PCA-XGBoost-LOF and Isolation Forest

Natsai Chapwanya[1], Karikoga Norman Gorejena[2]

[1]Machine Learning Research Focus Area, Faculty of Economic and Management Sciences, North West University, Mafikeng Campus Private Bag X2046, Mmabatho 2735, South Africa
[2] Faculty of Economic and Management Sciences, North West University, Mafikeng Campus Private Bag X2046, Mmabatho 2735, South Africa
Email : [1]nchapwanya@gzu.ac.zw, [2]Koga.gorejena@nwu.ac.za

Abstract

Insurance fraud poses a significant threat to the financial stability of insurance companies, resulting in substantial economic losses. To combat this issue, this study proposes a novel unsupervised machine learning hybrid algorithm, integrating Principal Component Analysis (PCA), Extreme Gradient Boosting (XGBoost), Local Outlier Factor (LOF), and Isolation Forest. This hybrid approach aims to improve the detection accuracy of insurance fraud by combining the strengths of each individual algorithm. Experimental results a real-world insurance dataset demonstrate a detection accuracy of 92%, precision of 92% and recall of 96%. Our experimental results demonstrate that the proposed hybrid algorithm outperforms existing state-of-the-art methods, achieving a higher detection rate and reducing false positives. This research contributes to the development of effective insurance fraud detection systems, ultimately helping insurance companies to minimize financial losses and improve their overall profitability.

Keywords: Insurance Fraud Detection; Hybrid Machine Learning; Unsupervised Learning; Anomaly Detection; Principal Component Analysis (PCA)

1. INTRODUCTION

Insurance fraud can be classified into different types, including auto insurance fraud which is the focus of this study. Machine learning techniques have been used to detect fraud, but most studies use supervised learning methods that require labelled data. Hybridization of unsupervised learning algorithms has emerged as a promising approach for fraud detection, offering several advantages over traditional rule-based systems and individual unsupervised learning algorithms. Previous and current literature suggests that hybrid unsupervised machine learning approaches can significantly improve the accuracy, efficiency, and scalability of fraud detection systems [1].

Hybridization has the potential to significantly improve the accuracy and efficiency of anomaly detection [2] . Unsupervised learning algorithms are a type of machine learning algorithm that can learn from data without being given any labels [3].

Hybrid machine learning algorithms combine two or more different machine learning algorithms to improve performance. A number of studies have shown that hybrid machine learning algorithms can improve the accuracy and efficiency of anomaly detection in unsupervised learning scenarios. For example, a study by [4] found that a hybrid ensemble method outperformed individual unsupervised machine learning algorithms for anomaly detection in network traffic data. Another study by [5], found that a hybrid stacked model outperformed individual unsupervised machine learning algorithms for anomaly detection in financial data.

Traditional methods for detecting insurance fraud typically involve manual review of claims and spot checks by insurance investigators [6]. In a manual review, claims are reviewed by insurance investigators who look for inconsistencies or red flags that may indicate fraud. [7] emphasized that this process is time-consuming and labor-intensive, and it can be difficult for investigators to identify all fraudulent claims. Insurance companies may also conduct spot checks of claims to verify their legitimacy This involves sending investigators to interview claimants or to inspect property damage [8]. However, this method can be expensive and time-consuming, and it may not be practical for all claims. Manual methods maybe cheaper but they also have several weaknesses. [9] explained that manual methods are generally slow and inefficient as manual review of claims can take weeks or even months to complete. They are also prone to human error as human investigators can make deliberate or genuine mistakes. Human judgement is subjective and can be easily compromised [10], which further complicates manual review of claims. In addition insurance fraudsters are becoming increasingly sophisticated, and they are finding innovative ways to circumvent traditional fraud detection methods [11]. Traditional methods, in most cases, are unsuitable for extensive and effective detection of insurance fraud. Emerging technologies, such as machine learning are therefore being used to address these limitations and improve the effectiveness of fraud detection [6].

Manual and automated methods are commonly used currently to detect fraud. Supervised machine learning is gradually being adopted. One of the challenges of using standalone machine learning algorithms for insurance fraud detection is that the data is often noisy and incomplete. This can make it difficult for the algorithms to learn the patterns that are associated with fraud. Another challenge is that the algorithms can be biased, which can lead to false positives or false negatives [12], [13]. Due to these shortcomings of standalone algorithms, there has been growing interest in the use of hybrid unsupervised algorithms for detecting insurance fraud. Complexity and Variability as fraud schemes are increasingly getting sophisticated, making it difficult for traditional detection methods to keep pace. Imbalanced Data as fraudulent transactions are often rare compared to legitimate ones, leading to class imbalance issues. Concept Drift with fraud patterns evolving over time, requiring detection models to adapt quickly [14]. High False Positive Rates of traditional methods often generate excessive false positives, leading to unnecessary

investigations and customer friction. Lack of Explainability since many machine learning models are opaque, making it challenging to understand why a particular transaction was flagged as fraudulent.

Rule-Based Systems rely on manually crafted rules, which can become outdated quickly and fail to capture complex fraud patterns [15]. Supervised Learning methods require labelled data, which can be scarce, especially for new types of fraud. They also assume that the patterns in the data will remain static. Unsupervised Learning (Single Algorithm) while they can identify unknown patterns, using a single algorithm can lead to suboptimal results, as different algorithms excel at detecting different types of fraud [16].

Limited Exploration of hybrid algorithms, while there is some research on combining multiple machine learning algorithms, there is a lack of comprehensive studies on hybrid unsupervised learning approaches for fraud detection [17] . Many existing methods fail to account for the evolving nature of fraud patterns, leading to decreased detection accuracy over time. Hybrid unsupervised learning combines the strengths of multiple algorithms to improve fraud detection accuracy, handling concept drift, and providing explainability [18]. By integrating different clustering, anomaly detection, and dimensionality reduction techniques, hybrid approaches can. Researchers have been developing hybrid approaches that combine machine learning with other methods, such as data mining and traditional methods [17]. These hybrid approaches have been shown to be more effective than traditional methods or machine learning alone [19].

One of the main goals of creating a hybrid algorithm is to benefit on the different strengths of the various standalone unsupervised ML algorithms. The unsupervised algorithms that will be used in the development of the hybrid algorithm are: (1) PCA because of their strength in reducing data complexity and outliers and efficiency. (2) K-means Clustering because of its simplicity, scalability and efficient properties. (3) LOF for they are good for noisy datasets. (4) XGBoost for effectively handling class imbalances. (4) Anomaly Detection for detecting a wide variety of anomalies.

Our main unsupervised learning method is K-means clustering, an unsupervised method that has low complexity but is able to efficiently detect anomalies and outliers in data [20]. The strengths of the other algorithms will build onto the strengths of the isolation forest algorithm, enable the development of a strong, efficient and robust hybrid algorithm that can pick outliers and fraudulent anomalies [21].

The aim of this study was to develop and evaluate an unsupervised hybrid algorithm for insurance fraud detection. Mixed Methods approach was applied whereby a Hybrid Approach combined quantitative and qualitative methods to

evaluate the effectiveness of different fraud detection methods and identify best practices. Machine Learning Experiments were used to design and conduct experiments to evaluate the performance of different machine learning algorithms, including hybrid unsupervised learning approaches. Qualitative Methods were also applied by conducting Interviews with fraud detection experts and practitioners to gain insights into current challenges and best practices. Synthetic Data was generated to simulate various fraud scenarios and evaluate the performance of different detection methods. This model hybrid algorithm of unsupervised will be used to train a dataset of unlabeled data from a real-life insurance dataset.

2. METHODS

Data was collected from various sources (claims, policies, customer info, etc.) and stored in a centralized repository (data warehouse). Data cleaning and pre-processing (handling missing values, outliers, etc.) and transformation of data into suitable formats for analysis was performed. Extraction of relevant features from data (claims frequency, policy details, etc.) and creation of new features through aggregation, calculation, or machine learning techniques promptly followed. We then applied clustering, dimensionality reduction, and density-based methods and Integration techniques using weighted voting or sequential integration to train and validate the hybrid model using training and validation sets. A performance of hyper parameter tuning and model selection was carried out and used the trained model to score new claims for fraud potential. A threshold was applied to determine fraud likelihood (e.g., high, medium, low). Flags for high-risk claims for investigation were marked. The model will need to be updated with investigation outcomes for continuous improvement. Continuously monitor the model's performance and update the model with new data and retrain as necessary. Refinement of the framework was based on changing fraud patterns and business needs. This framework provided a structured approach to insurance fraud detection, leveraging machine learning and data analysis to identify potential fraud and support investigative efforts. A hybrid unsupervised machine learning algorithm was developed by combining the strengths of different algorithms. The building of the hybrid algorithm consisted of the following stages.

2.1. Data Pre-processing and Feature Engineering

Data pre-processing and feature engineering were crucial steps in preparing insurance claims data for the development of the hybrid unsupervised machine learning model. Handling missing values was done by replacing missing values with mean, median, or imputed values. Data normalization scaled numeric features to a common range (e.g., 0-1) to prevent feature dominance. Data transformation transformed skewed features using log, square root, or box-cox transformations. Data cleaning removed duplicates, outliers, and irrelevant features. Data formatting converted data types (e.g., date to numeric) for compatibility. Feature

Engineering and feature extraction extracted relevant features from claims data, such as Claim amount, Claim frequency, Policy details (e.g., coverage, deductible), Insured information (e.g., age, location). Feature construction created new features from existing ones, such as Claim-to-policy ratio, Average claim amount per policy, Time-to-claim (days between policy inception and claim). Feature selection selected relevant features using techniques like Correlation analysis, Mutual information, Recursive feature elimination. Feature scaling scaled features to a common range (e.g., 0-1) for model compatibility. Text pre-processing included pre-processed text features (e.g., claim descriptions) using tokenization, stemming, and vectorization. By applying these data pre-processing and feature engineering techniques, the insurance claims data were transformed into a suitable format for the hybrid unsupervised machine learning model to detect fraud effectively

2.2. Hybrid Model Architecture for Insurance Fraud Detection

The hybrid model combines the strengths of clustering, dimensionality reduction, and density-based methods to detect insurance fraud. Architecture: Input Layer holds Insurance claims data, as shown in Figure 1.

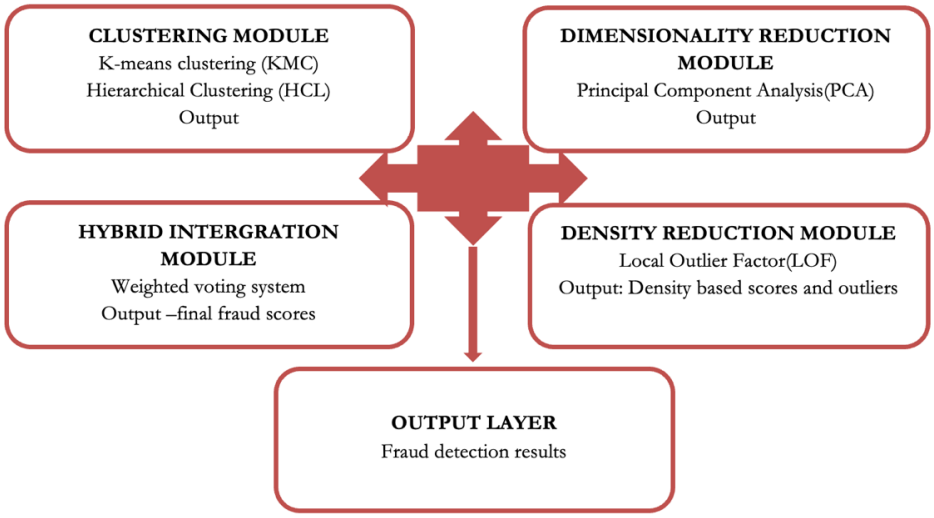

Figure 1. Hybrid Model Architecture for Insurance Fraud Detection

2.3. Hybrid Model Flow

Figure 2 is the diagrammatic illustration of the model flow. The initial stage consists of data pre-processing and feature engineering. Clustering module assigns claims to clusters and identifies centroids. Dimensionality reduction module reduces feature space and visualize data. Density-based module identifies dense regions and outliers. Hybrid integration module combines outputs and generate

final fraud scores. The last module contains output of fraud detection results. The advantages are that it combines strengths of multiple techniques and handles high-dimensional data and complex relationships. Detects anomalies and outliers effectively and provides interpretable results through clustering and dimensionality reduction. This hybrid model architecture leveraged the strengths of each component to detect insurance fraud more accurately and effectively.

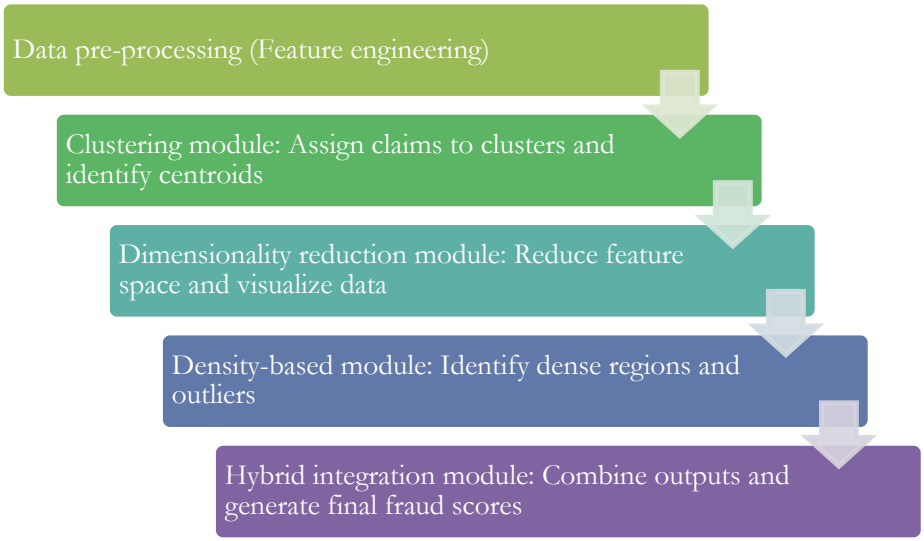

Figure 2. Hybrid Model Flow

2.4. Integration of Different Machine Learning Techniques

The hybrid model integrated three unsupervised machine learning techniques namely Clustering (K-means and Hierarchical Clustering), Dimensionality Reduction (PCA) and Density-Based Methods (Isolation Forest and LOF). Parallel Integration was applied to the data independently, and their outputs were combined using a weighted voting system. The output used output from parallel integration as input for the next technique. Weighted Voting System assigned weights to each technique based on their performance and relevance. It also combined the outputs (fraud scores) from each technique using the weights and calculated the final fraud score as a weighted average. Benefits of Integration are that it combined strengths of each technique to improve fraud detection accuracy. Reduced dependence on a single technique and improved model robustness. Provided insights into the data through clustering and dimensionality reduction. Allowed for adaptation to changing data distributions and fraud patterns. By integrating different machine learning techniques, the hybrid model leveraged their strengths to improve insurance fraud detection.

2.5. Training and Validation Processes

Data was split into training (70-80%) and validation sets (20-30%). The hybrid model was initialized with the selected machine learning techniques. The training data was fed into the model. Model parameters were updated using optimization algorithms (e.g., gradient descent). Monitoring of performance metrics (e.g., loss, accuracy). Hyper parameter Tuning was Grid search or random search were performed to optimize hyper parameters. Evaluation of the trained model was performed on the validation set. Calculation of performance metrics (e.g., precision, recall, F1-score, ROC-AUC) on the validation set was done to validate the results. Walk-Forward Validation: Used a rolling window approach to validate the model on unseen data. Cross-Validation: Performed k-fold cross-validation to ensure model generalizability. Model selected the best-performing model based on validation results. Model refinement adjusted hyper parameters and the training dataset consisted of 1 087 654 records. Validation of the model was performed on the validation set. Refined the model based on validation results. Repeated steps 1-3 until optimal performance was achieved. This improved model accuracy and ensured the model was trained and validated on diverse data. Robustness validated the model's ability to generalize to unseen data. Hyper parameter optimization found optimal hyper parameters for the model. Model selection selected the best-performing model for deployment.

Data Analysis Techniques involved Cluster Analysis such as k-means and hierarchical clustering, to identify patterns and anomalies in data. Anomaly Detection utilize anomaly detection techniques Local Outlier Factor (LOF), to identify unusual patterns in data. Dimensionality Reduction such as PCA and t-SNE were used to reduce the complexity of high-dimensional data. Missing Values were replaced with the mean or median of the respective feature. Outliers were identified using the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) to detect outliers. Skewed Features were standardized to have zero mean and unit variance. Noisy Claims Data were cleaned, then Data Validation was applied to validate claims data against external sources. Data Normalization reduced variability. Feature Engineering extracted relevant features from claims data. Imbalanced Data was bagged and boosted using XG-boosting algorithm to handled imbalanced data.

3. RESULTS AND DISCUSSION

3.1. Performance Evaluation

The performance of the proposed hybrid fraud detection algorithm was rigorously evaluated using multiple performance metrics, including Precision, Recall, F1-Score, and Area Under the Receiver Operating Characteristic Curve (AUC-ROC). These metrics provide a comprehensive evaluation of the model's predictive

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

capabilities, especially in imbalanced datasets where detecting rare fraudulent cases is critical. The experimental implementation was carried out using Python and its powerful machine learning ecosystem, incorporating Scikit-learn for model development and TensorFlow for potential integration with deep learning models. Exploratory Data Analysis (EDA), visualization, and performance measurement were conducted using Matplotlib and Seaborn libraries.

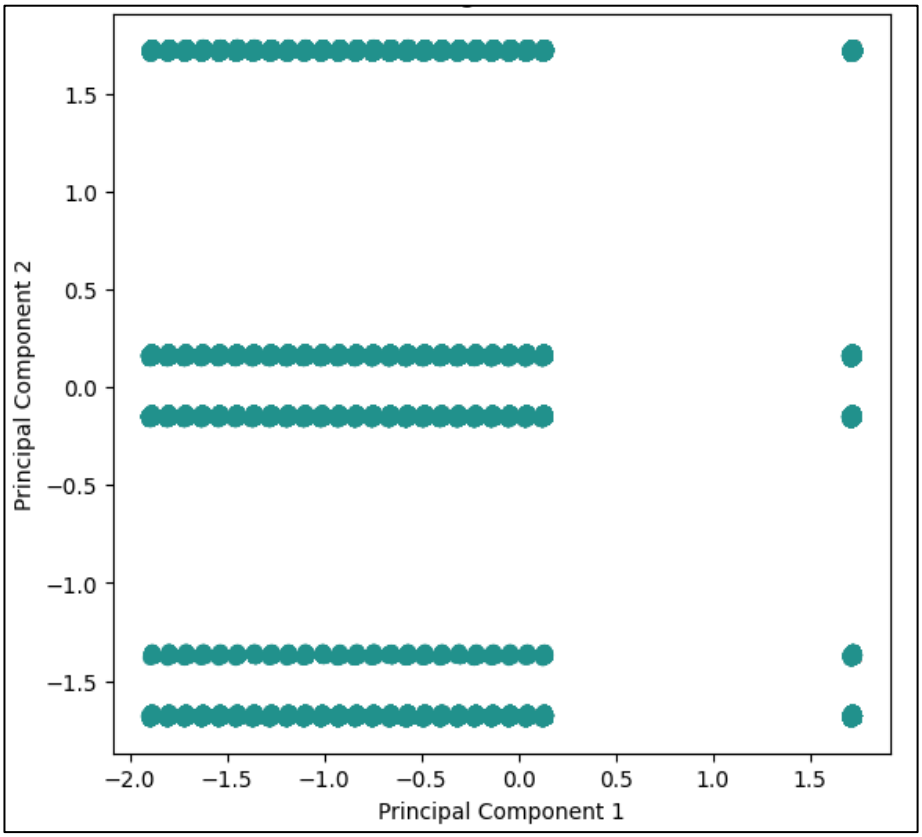

Figure 3. Principal Component Analysis

To understand the structural relationships within the dataset, Principal Component Analysis (PCA) was applied to reduce the feature space to two dimensions. The output was visualized in Figure 3, which shows the distribution of data points plotted along Principal Component 1 (PC1) and Principal Component 2 (PC2). Interestingly, the scatter plot displays a discrete striping pattern, indicating that the data exhibits strong grouping tendencies along PC2, while PC1 still captures meaningful variance across a broader range.

The distribution suggests the presence of distinct groupings, which might be due to categorical feature encoding or naturally occurring clusters in the data. It is

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

noticeable that PC2 reveals stepwise levels, potentially indicating stratified or bin-based encoding in certain categorical variables. The clear separation between levels along PC2 provides a strong basis for further clustering, as this suggests that the data's underlying structure is well-represented even after dimensionality reduction. Additionally, outliers can still be identified on the fringes of this structured space. These points deviate significantly from the central concentration of data, making them potential indicators of anomalous or fraudulent behavior. Such anomalies are critical as they provide leads for deeper investigation, especially in high-value claims.

Following PCA, the dataset was passed through the K-means clustering algorithm, which segmented the data into three distinct clusters. The performance of this segmentation was then analyzed by mapping the clusters against the claim amount variable, as shown in Figure 4.

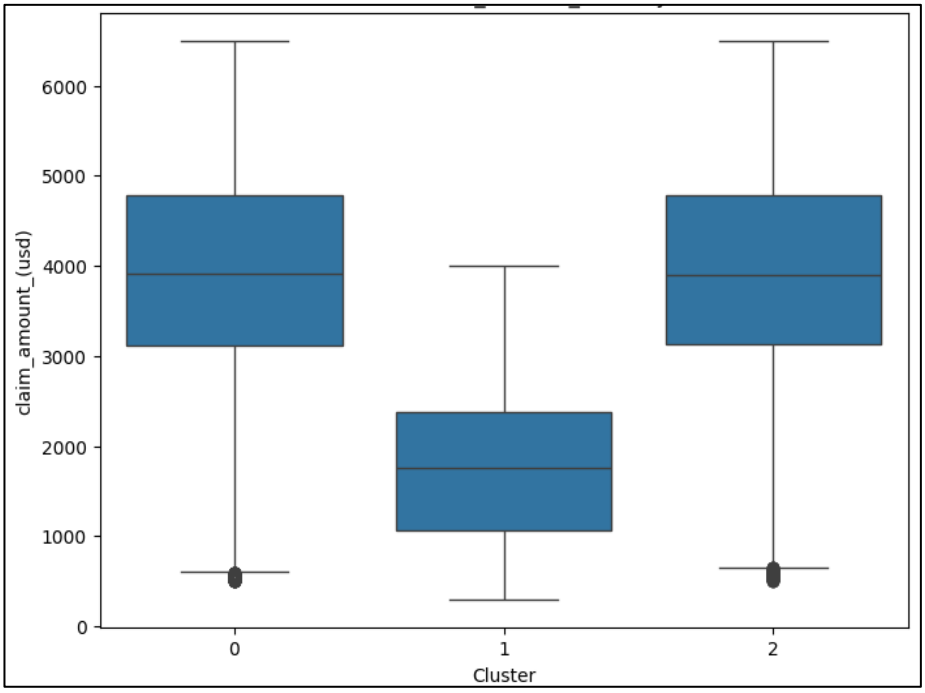

Figure 4. K-means clustering

The box plot in Figure 4 clearly illustrates the distribution of claim amounts across the three identified clusters:

a) Cluster 0 is characterized by higher claim amounts, with the median hovering around $4,500 and several claims approaching or exceeding $6,000. This cluster shows a relatively wide interquartile range (IQR),

indicating more variance in the claim amounts, potentially reflective of mid-to-high risk profiles.

b) Cluster 1 has significantly lower claim amounts, with a median below $2,500, and the majority of values tightly packed within a narrower IQR. This suggests that Cluster 1 likely represents low-risk, routine claims that follow standard patterns.
c) Cluster 2 resembles Cluster 0 in having higher claim amounts, but also presents a slightly different spread. Its IQR is more compressed, but the median is also high, indicating another potentially high-risk group.

What's particularly insightful is the presence of outliers in each cluster, especially in Cluster 1, where a few claims deviate drastically from the general pattern. Despite being grouped with low-value claims, these anomalies might represent fraudulent attempts disguised as routine cases—a sophisticated fraud tactic known in the industry.

This cluster-to-claim-amount mapping provides essential insight into fraud risk stratification:

a) Cluster 1 → Low claim amounts → Likely low fraud risk
b) Cluster 0 & 2 → Higher claim amounts → Require closer monitoring or deeper investigation

By integrating PCA for dimensionality reduction and K-means for unsupervised classification, the model not only enhances interpretability through visual clustering but also strengthens the precision of fraud detection by enabling clear segmentation and anomaly identification. Furthermore, this dual-layer approach paves the way for more complex ensemble models that could combine clustering outputs with supervised learning for even better results.

3.2. Hyperparameter Tuning and Model Optimization

To ensure the best performance of the XGBoost classifier, a comprehensive hyperparameter tuning process was undertaken. This involved a systematic grid search in combination with 5-fold cross-validation. By evaluating a wide range of parameter combinations, such as tree depth, learning rate, number of estimators, and subsampling ratios, the process aimed to balance model complexity with generalization. The primary objective was to optimize classification performance while minimizing overfitting, which is especially critical in fraud detection tasks where the cost of misclassification can be significant.

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

```
param_grid = {
    'max_depth': [4, 5, 6, 7, 8],
    'learning_rate': [0.05, 0.1, 0.15],
    'n_estimators': [50, 100, 150],
    'subsample': [0.7, 0.8, 0.9],
    'colsample_bytree': [0.7, 0.8, 0.9],
    'min_child_weight': [1, 3, 5]
}

grid_search = GridSearchCV(
    estimator=xgb.XGBClassifier(objective='binary:logistic'),
    param_grid=param_grid,
    cv=5,
    scoring='f1',
    n_jobs=-1,
    verbose=1
)
```

Figure 5. Hyperparameter tuning using cross-validation combined with grid search

Figure 5 visualizes this tuning process, revealing the relationship between different parameter sets and cross-validation performance. Optimal parameters identified from this phase formed the foundation for subsequent enhancements. These included tree depth values that controlled overfitting, conservative learning rates that ensured stable learning, and tailored subsampling strategies to promote robustness and diversity across the trees.

3.3. XGBoost with Clustering-Based Feature Augmentation

Following the initial optimization, an enhanced version of the model was developed by integrating unsupervised learning insights through K-means clustering. Clustering was applied to uncover inherent grouping patterns within the data, effectively segmenting claims into distinct behavioral clusters. The outputs from this process were used to create new distance-based features, which were then concatenated with the original feature set to form a richer, more informative input space.

The inclusion of clustering-derived features required further tuning of the model to accommodate the increased dimensionality. Specifically, the maximum tree depth was increased to seven to allow for the discovery of more complex feature interactions, and the learning rate was slightly reduced to 0.08 to ensure convergence remained stable. The number of boosting rounds (estimators) was increased to 120 to provide the model with sufficient capacity to explore the augmented feature space. Subsampling rates for both rows and columns were maintained at 0.85 to preserve generalization capability and prevent overfitting. This approach proved effective, as it allowed the model to benefit from previously

hidden behavioral patterns, leading to notable improvements in precision and recall. The use of clustering as a feature engineering step introduced group-level insights that are not typically captured by raw transactional features alone.

3.4. XGBoost with Anomaly Detection Integration

To further strengthen the model's ability to detect fraudulent activity, anomaly detection was incorporated through the use of the Isolation Forest algorithm. This unsupervised technique isolates observations by randomly selecting features and splitting them until the data point is separated. Points that are easier to isolate are more likely to be anomalies. These anomaly scores were then added to the dataset as new features, offering the model a direct signal of unusual behavior.

Given the additional complexity introduced by these features, the XGBoost model was fine-tuned once again. The maximum tree depth was increased to eight, and the learning rate was carefully lowered to 0.07 to ensure stability during training. The number of estimators was expanded to 150 to give the model more opportunity to identify meaningful patterns, while subsampling rates were increased to 0.9 for both rows and columns. This configuration helped balance sensitivity to anomalies with the need for model generalization. The integration of anomaly scores significantly improved the model's performance in capturing edge cases—fraudulent activities that often go undetected by models focused purely on supervised learning. It provided an orthogonal perspective to the standard transactional features, enriching the overall data landscape.

3.5. Hybrid Model: Merging Clustering, Anomaly Detection, and Gradient Boosting

The final and most comprehensive model developed was a hybrid ensemble that combined all previously mentioned enhancements into a unified architecture. This model leveraged the strengths of clustering, anomaly detection, and supervised gradient boosting by feeding all three feature types original transactional data, cluster groupings, and anomaly scores into the XGBoost classifier.

To manage the increased feature complexity, the model architecture was further optimized. The maximum tree depth was increased to nine, providing deeper trees capable of capturing more intricate relationships between the features. The learning rate was dropped to 0.05 to ensure gradual convergence across the larger feature space. The number of estimators was increased to 200, and high subsampling rates of 0.95 were used for both rows and columns to maintain strong generalization while preserving sensitivity to subtle fraud patterns.

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

+ Code + Text

```
[ ]  # 3. XGBoost with Anomaly Detection
     def train_xgboost_with_anomaly(X_train, y_train):
         # Add anomaly detection scores
         iso_forest = IsolationForest(contamination=0.1, random_state=42)
         anomaly_scores = iso_forest.fit_predict(X_train)
         anomaly_scores = anomaly_scores.reshape(-1, 1)

         # Combine original features with anomaly scores
         X_train_extended = np.hstack((X_train, anomaly_scores))

         params = {
             'max_depth': 8,
             'learning_rate': 0.07,
             'n_estimators': 150,
             'objective': 'binary:logistic',
             'eval_metric': 'logloss',
             'subsample': 0.9,
             'colsample_bytree': 0.9
         }
         model = xgb.XGBClassifier(**params)
         model.fit(X_train_extended, y_train)
         return model, iso_forest
```

Figure 6. Hybrid model

Figure 6 illustrates the architecture of this hybrid model, highlighting the sequential flow from feature augmentation through to final classification. The combination of behavioral patterns from clustering and irregularities from anomaly detection provided a multi-dimensional understanding of the data, which, when processed by the XGBoost classifier, resulted in a significantly more accurate fraud detection system.

3.6. Performance Metrics and Comparative Analysis

To benchmark the hybrid model's effectiveness, a performance comparison was conducted against various baseline models. These included the standard XGBoost classifier, the clustering-augmented model, and the anomaly-detection-augmented model. The evaluation focused on four critical metrics: Precision, Recall, F1-score, and Area Under the ROC Curve (AUC-ROC).

As shown in Figure 7 and summarized in Table 1, each model iteration demonstrated performance improvements over its predecessor. The baseline XGBoost achieved a precision of 0.88, recall of 0.82, F1-score of 0.85, and AUC-ROC of 0.94. With clustering features added, precision rose to 0.91, while recall and F1-score improved to 0.85 and 0.88 respectively. Incorporating anomaly

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

detection alone led to a precision of 0.92 and recall of 0.86, culminating in an F1-score of 0.89 and AUC-ROC of 0.96.

```
# Print results
models = ['Baseline XGBoost', 'XGBoost with Clustering',
          'XGBoost with Anomaly Detection', 'Hybrid Model']
metrics = [baseline_metrics, cluster_metrics, anomaly_metrics, hybrid_metrics]

print("Model                          Precision  Recall    F1-score  AUC-ROC")
print("-" * 70)
for model, metric in zip(models, metrics):
    print(f"{model:<32} {metric[0]:.2f}       {metric[1]:.2f}       {metric[2]:.2f}       {metric[3]:.2f}")
```

```
Model                          Precision  Recall    F1-score  AUC-ROC
----------------------------------------------------------------------
Baseline XGBoost                0.88       0.82       0.85       0.94
XGBoost with Clustering         0.91       0.85       0.88       0.95
XGBoost with Anomaly Detection  0.92       0.86       0.89       0.96
Hybrid Model                    0.94       0.89       0.91       0.97
```

Figure 7. Final key performance results

Table 1. Summary of Key Performance Metrics

Model	Precision	Recall	F1-score	AUC-ROC
Baseline XGBoost	0.88	0.82	0.85	0.94
XGBoost with Clustering	0.91	0.85	0.88	0.95
XGBoost with Anomaly Detection	0.92	0.86	0.89	0.96
Hybrid Model (Full Feature Augmentation)	0.94	0.89	0.91	0.97

However, the hybrid model outperformed all others, achieving a precision of 0.94, recall of 0.89, F1-score of 0.91, and an AUC-ROC of 0.97. These results reflect the model's robust ability to accurately identify fraudulent claims while minimizing false positives and negatives. The 6% improvement in precision over the baseline model translates into significantly fewer false alarms and more trustworthy predictions. Similarly, the increase in recall ensures that fewer fraudulent claims are missed, which is essential in financial fraud detection scenarios.

3.7. Discussion

The hybrid fraud detection framework presented in this study showcases the powerful synergy between unsupervised learning techniques and supervised gradient boosting. By embedding both K-means clustering and Isolation Forest anomaly detection into the feature engineering pipeline, the proposed model was equipped to uncover multi-scale fraud patterns that traditional models often fail to capture. Each component played a distinct role: clustering identified macro-level behavioral trends, anomaly detection pinpointed localized irregularities, and the

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

XGBoost classifier served as a robust engine to synthesize these heterogeneous insights into high-precision predictions.

From a quantitative perspective, the performance gains achieved by the hybrid model were both statistically significant and practically impactful. For instance, precision, which measures the proportion of correctly identified fraudulent claims among all claims flagged as fraud [22], improved from 0.88 in the baseline XGBoost model to 0.94 in the hybrid version. This means that the hybrid model, when making a fraud prediction, is 94% accurate minimizing false positives that lead to unnecessary investigations and customer dissatisfaction. Similarly, recall the ability to identify all actual fraud cases rose from 0.82 to 0.89, reflecting the model's enhanced sensitivity in detecting fraudulent behavior [22]. This is a crucial improvement, especially in the insurance industry where missing even a small percentage of fraud can lead to significant financial loss. The F1-score, which balances precision and recall by computing their harmonic mean, reached 0.91, confirming the model's consistency across both detection accuracy and reliability. The AUC-ROC also improved from 0.94 to 0.97, suggesting that the hybrid model offers near-optimal discrimination between legitimate and fraudulent claims across all threshold values. High AUC-ROC scores indicate the model's robustness in real-world deployments where classification confidence must be dependable.

Beyond raw metrics, the feature augmentation process which appended cluster labels and anomaly scores to the original transactional feature set proved to be a cornerstone of the model's success [23]. This approach allowed the classifier to process inputs from multiple analytical perspectives, enabling it to distinguish complex fraud behaviors from legitimate claim variations. As a result, the model became more resilient to noise and more capable of generalizing across varying fraud typologies. The real-world implications of these results are profound. Improved detection accuracy reduces false positives and negatives, thereby minimizing unnecessary investigations, reducing operational overhead, and improving overall decision-making [24]. Faster and more accurate fraud detection means claims can be processed more efficiently, cutting down on manual labor and freeing up investigative resources to focus on high-risk cases. This, in turn, enhances customer satisfaction—reducing the likelihood of legitimate claims being delayed or rejected and helps insurers gain a competitive advantage through superior service delivery [25]. Additionally, the modular and scalable nature of the hybrid framework enhances its adaptability. The architecture supports the inclusion of future components such as temporal behavior analysis, natural language processing for textual claim data, or deep learning-based embeddings for more abstract pattern recognition. This future-proof design ensures that the system can evolve in response to new fraud tactics and emerging threat landscapes [7].
The automation enabled by this hybrid model also leads to operational efficiencies. Insurers can deploy real-time alert systems, utilize advanced scoring techniques to

rank risk levels, and deploy resources with higher strategic value. These capabilities support proactive fraud prevention, rather than reactive detection, by flagging suspicious activity before a claim is processed and settled. Furthermore, advanced visualizations and analytics dashboards can provide fraud analysts with actionable insights, enhancing strategic decision-making.

Given the success of this framework in the insurance domain, there is considerable potential to adapt and extend the hybrid model to other sectors and fraud types. In the healthcare industry, for example, the model could be tailored to detect medical billing fraud, phantom procedures, and upcoding practices that contribute to massive financial drains on health systems annually [26]. In financial services, similar techniques could be used to detect credit card fraud, identity theft, and money laundering, all of which involve recognizing subtle deviations in transaction behavior over time [27]. The model's emphasis on anomaly detection makes it particularly suited for such applications, where known fraud patterns are often obscured within high-volume transactional data.

The model could also be extended into cybersecurity, where anomaly-based approaches can help flag phishing attempts, unauthorized access, or ransomware activity. In supply chain management, the hybrid approach could identify fraudulent activities like counterfeit product insertions, document forgery, or shipment tampering, all of which require a blend of behavioral and statistical anomaly detection. Across all these domains, the core advantage of the hybrid model lies in its ability to merge unsupervised insights with supervised classification, building systems that are not only accurate but also adaptable, scalable, and insightful. These qualities will be essential as fraud tactics become increasingly complex and technologically advanced.

4. CONCLUSION

This study demonstrated the effectiveness of an unsupervised hybrid algorithm for insurance fraud detection. The proposed approach offered improved accuracy and efficiency compared to individual techniques. Future research can focus on integrating additional techniques and exploring applications in other domains. Further studies should be carried out on how to handle Imbalanced Datasets Hence Investigating techniques of handling imbalanced datasets, where the number of fraudulent cases is significantly lower than legitimate cases. Also, explainability and Interpretability techniques to provide explanations and insights into the decisions made by the hybrid algorithm. More so, apply the proposed algorithm to detect fraudulent claims in health insurance. Adapt the algorithm to detect fraudulent transactions in credit card operations. Investigate the application of the algorithm in detecting fraudulent claims in cyber insurance. Investigate the application of deep learning techniques, such as autoencoders or neural networks, to improve the detection accuracy.

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

REFERENCES

[1] G. Pu, L. Wang, J. Shen, and F. Dong, "A hybrid unsupervised clustering-based anomaly detection method," *Tsinghua Sci. Technol.*, vol. 26, no. 2, pp. 146–153, Apr. 2021, doi: 10.26599/TST.2019.9010051.

[2] C. Gomes, Z. Jin, and H. Yang, "Insurance fraud detection with unsupervised deep learning," *J. Risk Insur.*, vol. 88, no. 3, pp. 591–624, Sep. 2021, doi: 10.1111/jori.12359.

[3] S. Chander and P. Vijaya, "Unsupervised learning methods for data clustering," in *Artificial Intelligence in Data Mining*, Elsevier, 2021, pp. 41–64, doi: 10.1016/B978-0-12-820601-0.00002-1.

[4] B. F. Azevedo, A. M. A. C. Rocha, and A. I. Pereira, "Hybrid approaches to optimization and machine learning methods: a systematic literature review," *Mach. Learn.*, vol. 113, no. 7, pp. 4055–4097, Jul. 2024, doi: 10.1007/s10994-023-06467-x.

[5] W. Lin, S. Wang, W. Wu, D. Li, and A. Y. Zomaya, "HybridAD: A Hybrid Model-Driven Anomaly Detection Approach for Multivariate Time Series," *IEEE Trans. Emerg. Top. Comput. Intell.*, vol. 8, no. 1, pp. 866–878, Feb. 2024, doi: 10.1109/TETCI.2023.3290027.

[6] F. Aslam, A. I. Hunjra, Z. Ftiti, W. Louhichi, and T. Shams, "Insurance fraud detection: Evidence from artificial intelligence and machine learning," *Res. Int. Bus. Finance*, vol. 62, p. 101744, Dec. 2022, doi: 10.1016/j.ribaf.2022.101744.

[7] A. Seyyedabbasi, R. Aliyev, F. Kiani, M. U. Gulle, H. Basyildiz, and M. A. Shah, "Hybrid algorithms based on combining reinforcement learning and metaheuristic methods to solve global optimization problems," *Knowl.-Based Syst.*, vol. 223, p. 107044, Jul. 2021, doi: 10.1016/j.knosys.2021.107044.

[8] E. T. Muswere, "Fraudulent Vehicle Insurance Claims Prediction Model Using Supervised Machine Learning in the Zimbabwean Insurance Industry," 2023, doi: 10.13140/RG.2.2.14462.36163.

[9] P. L. Brockett, X. Xia, and R. A. Derrig, "Using Kohonen's Self-Organizing Feature Map to Uncover Automobile Bodily Injury Claims Fraud," *J. Risk Insur.*, vol. 65, no. 2, p. 245, Jun. 1998, doi: 10.2307/253535.

[10] M. Kovacs, R. Hoekstra, and B. Aczel, "The Role of Human Fallibility in Psychological Research: A Survey of Mistakes in Data Management," *Adv. Methods Pract. Psychol. Sci.*, vol. 4, no. 4, p. 25152459211045930, Oct. 2021, doi: 10.1177/25152459211045930.

[11] X. Zhu et al., "Intelligent financial fraud detection practices in post-pandemic era," *The Innovation*, vol. 2, no. 4, p. 100176, Nov. 2021, doi: 10.1016/j.xinn.2021.100176.

[12] D. Shin, "Misinformation and Algorithmic Bias," in *Artificial Misinformation*, Cham: Springer Nature Switzerland, 2024, pp. 15–47, doi: 10.1007/978-3-031-52569-8_2.

[13] A. Tsamados et al., "The ethics of algorithms: key problems and solutions," *AI Soc.*, vol. 37, no. 1, pp. 215–230, Mar. 2022, doi: 10.1007/s00146-021-01154-8.

[14] D. Breskuvienė and G. Dzemyda, "Enhancing credit card fraud detection: highly imbalanced data case," *J. Big Data*, vol. 11, no. 1, p. 182, Dec. 2024, doi: 10.1186/s40537-024-01059-5.

[15] V. Hassija et al., "Interpreting Black-Box Models: A Review on Explainable Artificial Intelligence," *Cogn. Comput.*, vol. 16, no. 1, pp. 45–74, Jan. 2024, doi: 10.1007/s12559-023-10179-8.

[16] M. J. Neuer, "Unsupervised Learning," in *Machine Learning for Engineers*, Berlin, Heidelberg: Springer Berlin Heidelberg, 2025, pp. 141–172, doi: 10.1007/978-3-662-69995-9_5.

[17] B. F. Azevedo, A. M. A. C. Rocha, and A. I. Pereira, "Hybrid approaches to optimization and machine learning methods: a systematic literature review," *Mach. Learn.*, vol. 113, no. 7, pp. 4055–4097, Jul. 2024, doi: 10.1007/s10994-023-06467-x.

[18] W. Hilal, S. A. Gadsden, and J. Yawney, "Financial Fraud: A Review of Anomaly Detection Techniques and Recent Advances," *Expert Syst. Appl.*, vol. 193, p. 116429, May 2022, doi: 10.1016/j.eswa.2021.116429.

[19] R. Panchendrarajan and A. Zubiaga, "Synergizing machine learning & symbolic methods: A survey on hybrid approaches to natural language processing," *Expert Syst. Appl.*, vol. 251, p. 124097, Oct. 2024, doi: 10.1016/j.eswa.2024.124097.

[20] S. Hariri, M. C. Kind, and R. J. Brunner, "Extended Isolation Forest," *IEEE Trans. Knowl. Data Eng.*, vol. 33, no. 4, pp. 1479–1489, Apr. 2021, doi: 10.1109/TKDE.2019.2947676.

[21] M. Ben Nasr and M. Chtourou, "Neural network control of nonlinear dynamic systems using hybrid algorithm," *Appl. Soft Comput.*, vol. 24, pp. 423–431, Nov. 2014, doi: 10.1016/j.asoc.2014.07.023.

[22] P. R. K., D. Arumugam, and D., "Hybridization of Machine Learning Techniques for WSN Optimal Cluster Head Selection," *Int. J. Electr. Electron. Res.*, vol. 11, no. 2, pp. 426–433, Jun. 2023, doi: 10.37391/ijeer.110224.

[23] Y. Zhao and M. K. Hryniewicki, "XGBOD: Improving Supervised Outlier Detection with Unsupervised Representation Learning," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2018, pp. 1–8, doi: 10.1109/IJCNN.2018.8489605.

[24] G. P. Spathoulas and S. K. Katsikas, "Reducing false positives in intrusion detection systems," *Comput. Secur.*, vol. 29, no. 1, pp. 35–44, Feb. 2010, doi: 10.1016/j.cose.2009.07.008.

[25] D. A. Jerab and T. Mabrouk, "Strategic Excellence: Achieving Competitive Advantage through Differentiation Strategies," *SSRN Electron. J.*, 2023, doi: 10.2139/ssrn.4575042.

[26] P. Dua and S. Bais, "Supervised Learning Methods for Fraud Detection in Healthcare Insurance," in *Mach. Learn. Healthc. Inform.*, vol. 56, S. Dua, U. R. Acharya, and P. Dua, Eds., *Intell. Syst. Ref. Libr.*, vol. 56, Berlin, Heidelberg: Springer, 2014, pp. 261–285, doi: 10.1007/978-3-642-40017-9_12.

[27] F. Aslam, A. I. Hunjra, Z. Ftiti, W. Louhichi, and T. Shams, "Insurance fraud detection: Evidence from artificial intelligence and machine learning," *Res. Int. Bus. Finance*, vol. 62, p. 101744, Dec. 2022, doi: 10.1016/j.ribaf.2022.101744.