IOSR Journal of Business and Management (IOSR-JBM)
e-ISSN:2278-487X, p-ISSN: 2319-7668. Volume 27, Issue 3. Ser. 10 (March. 2025), PP 10-23
www.iosrjournals.org

Beyond Accuracy Metrics: Advanced Machine Learning Techniques for Effective Insurance Fraud Classification

Oluwaseun Oluwasola Mustapha[1],Oluwakemi Farinde[1] and Yakub Oufadi[2] *

Abstract
Insurance fraud occurs when an insured in- dividual, claimant, or company intentionally makes a false or misleading claim to gain fi- nancial benefit. Insurance applicants, policy- holders, third-party claimants, or profession- als from insurance companies or agencies can perpetrate insurance fraud at various stages of the insurance life-cycle. This study examines the efficiency and validity of popular machine learning algorithms for fraud prediction. In recent years, product recommendation, med- ical diagnosis, and image detection, among other fields, have greatly benefited from cur- rent machine learning algorithms, which have contributed to advancements in medicine and public safety. In this paper, we perform an analysis using five different algorithms: Ran- dom Forest (RF), Decision Tree (DT), Ad- aBoost, K-Nearest Neighbor (KNN), and XG- Boost to detect insurance fraud. Our findings indicate that DT provides the highest accuracy (79%) compared to the other techniques. The results of this study can assist risk analysts and professionals in evaluating the strengths and weaknesses of each model and in developing empirically effective decision rules to assess future insurance policies.

Date of Submission: 14-03-2025 | Date of Acceptance: 27-03-2025

I. Introduction

Due to the substantial amount of informa- tion available in the current era, many peo- ple are beginning to invest in various insur- ance policies. With increasing participation

*[1]Department of Mathematical Sciences, University of Essex, CO4 3SQ Colchester, United Kingdom (e- mail: om21222@essex.ac.uk).
[2]Space Science Center and Department of Physics, Uni- versity of New Hampshire, Durham, NH 03284, USA

in insurance policies comes a rise in insurance fraud. Insurance fraud is an intentional de- ception conducted against or on behalf of an insurance business or agent to obtain finan- cial gain. Over the years, a significant amount of money has moved across the insurance in- dustry as it is quite profitable. Insurance fraud also accounts for a substantial portion of costs incurred by insurance companies be- cause it reduces their earnings and has a long- term impact on their pricing policies. Several million dollars are misappropriated each year through insurance fraud. For example, in 2017, the Australian Insurance Fraud Bureau iden- tified false claims worth 280 *million*Insurance Fraud Bureau of Australia (2021). Addition- ally, 44,814 fraud claims totaling EUR 214 mil- lion were discovered in France in 2013, with the amount increasing to EUR 500 million in 2018 French Agency for the Fight against Insur- ance Fraud (2021). Statistics from the United States Coalition Against Insurance Fraud show that insurance fraud represents 17% of the to- tal compensation paid out by insurance com- panies, with an estimated annual value of 80 *billion* S.Jordon. (2016). According to the Insurance Bureau of Canada (IBC), car in- surance fraud totaled more than 542 million Canadian dollars in 2007. Chinese insurance officials reported that an average of RMB 35 billion was lost annually to insurance fraud in 2011, accounting for approximately 20% of all insurance company payments. Finally, it is es- timated that insurance fraud costs developing nations 600 *million* annually. Therefore, in- surance fraud is an international problem that harms nations and communities.

II. Literature Review

Since the development of artificial intelligence theory, machine learning techniques have been widely used for fraud detection. Itri, Mo- hamed, Mohammed, and Omar (2019) devel- oped a novel strategy to increase fraud pre- diction accuracy. Ten machine learning fraud prediction algorithms were evaluated for effec- tiveness and confirmability using auto insur- ance claim data. The study showed that Ran- dom Forest outperformed all other algorithms in predicting fraud. Machine learning ap- proaches are based on an analytical paradigm known as inductive reasoning, which derives conclusions from data patterns without mak- ing assumptions about functional forms such as probability distributions or linearity. Peng (2020) discuss how this flexibility necessitates extra caution to control the models' balance between generalization ability and complexity, as different models or even minor changes to their hyperparameters can have significant ef- fects on predictive performance when applied to the same dataset. Fraud detection is one of the main applications of machine learning in business administration and finance. From a decision-maker's perspective, this topic is ex- tremely relevant because decision support sys- tems that assist risk analysts in predicting fraudulent behavior directly affect a company's financial performance. This topic has been investigated by numerous researchers in re- cent years, as discussed in papers by Awoyemi, Adetunmbi, and Oluwadare (2017), Ngai, Hu, Wong, Chen, and Sun (2011), Raghavan and El Gayar (2019), and Waghade and Karandikar (2018). Sheshasaayee and Thomas (2018) dis- cuss the benefits of using machine learning methods to perform such tasks, especially con- cerning the most salient features of fraud- sters, while illustrating the main challenges faced by risk and fraud analysts in develop- ing fraud identification mechanisms and de- cision rules, given that the presence of fraud entails significant profit losses for the insur- ance sector. Similarly, Dal Pozzolo, Caelen, Le Borgne, Waterschoot, and Bontempi (2014) discussed the complexity involved in develop- ing a data-driven fraud detection algorithm, highlighting typical issues such as highly unbalanced class distributions, non-stationary distri- bution of data, lack of readily accessible micro- data due to confidentiality concerns, and on- going massive flows of new transactions. To address these issues, the authors assessed the predictive performance of three machine learn- ing models (Random Forest, Support Vector Machine, and Neural Network) using a dataset derived from actual credit card transactions. They also examined the overall effects of up- date periodicity, the use of balancing tech- niques, and the retention of older observations in the training dataset. For all training ap- proaches, the results showed that the Ran- dom Forest model consistently outperformed Support Vector Machines and Neural Net- works. Additionally, models that were updated with fresh data more frequently performed bet- ter, suggesting that fraud distributions can quickly change over time. Regarding the prob- lem of imbalanced classes, balancing methods were applied to improve performance over the "static" non-balanced dataset, in which Ran- dom Forest produced the worst performance. Finally, compared to keeping the dataset bal- anced, the method of removing earlier obser- vations showed marginally less benefit. Wang and Xu (2018) used Support Vector Machine (SVM), Random Forest, and Deep Neural Net- works to analyze descriptions of auto accidents to forecast frauds for auto insurance claims. All three models achieved an F1 Score greater than 75Conversely, Eshghi and Kargari (2019) argued that unsupervised methods like clus- tering and outlier detection may not be suf- ficient for complex fraud detection tasks and suggested a framework with Multi-Criteria De- cision Analysis and intuitionistic fuzzy sets to account for behavioral uncertainties when modeling the likelihood of fraudulent banking transactions. Similarly, Carcillo et al. (2021) advocated for combining supervised and un- supervised learning techniques for credit card fraud detection to more effectively adapt to changes in consumer behavior and fraudsters' ability to create novel fraud patterns. Based on clustering analysis, the authors developed outlier scores for various granularities, applied them to a real-world dataset, and reported improved detection effectiveness. In a recent study, Kim, Baik, and Cho (2016) proposed a multi-class algorithm to detect fraud inten- tion in financial misstatements using MetaCost Domingos1999 to incorporate asymmetric mis- classification costs and control for class imbal- ance; and Varmedja, Karanovic, Sladojevic, Arsenovic, and Anderla (2019) used SMOTE (Synthetic Minority Oversampling Technique) to balance training data along with Logistic Regression, Random Forest, Naive Bayes, and Neural Network as machine learning classi- fiers. Xu, Wang, Zhang, and Yang (2011) sug- gested a random rough subspace neural net- work ensemble-based method for detecting in- surance fraud. This method begins with a crude set reduction to create a set of reduc- tions that maintain stable data information. Next, a subset of reductions is assembled by selecting reductions randomly. Then, a neu- ral network classifier is trained using each of the chosen reductions on the insurance data. The trained neural network classifiers are sub- sequently combined using ensemble techniques. The effectiveness and efficiency of the pro- posed strategy were examined using a real- world vehicle insurance scenario. The findings demonstrate that a random rough subspace- dependent neural network ensemble technique can detect fraudulent insurance claims more quickly and accurately, making it a viable tool for detecting insurance fraud. Therefore, it is important to examine which machine learning models can more effectively identify fraud trends and precisely anticipate future offenses using real-world data and machine learning techniques. Given the wide range of fraud types that can occur, each with its own char- acteristics and method of operation Gottschalk (2010), in this paper we focus on frauds where the consumer is the perpetrator, specifically in policy claims of automobile insurance. The fact that the data for this study were obtained from a significant insurance provider enhances our understanding of the relative importance of the database elements, which is crucial for evaluating insurance policies in real-world con- texts.

III. PROPOSED MODEL (Data Collection, Exploration Pre- Processing)

Data overview, pre-processing, model applica- tion, and performance evaluation are the pri- mary steps of our proposed model. Each phase is crucial and enhances its effectiveness. The proposed model for detecting insurance fraud in this work is shown in Figure 1.

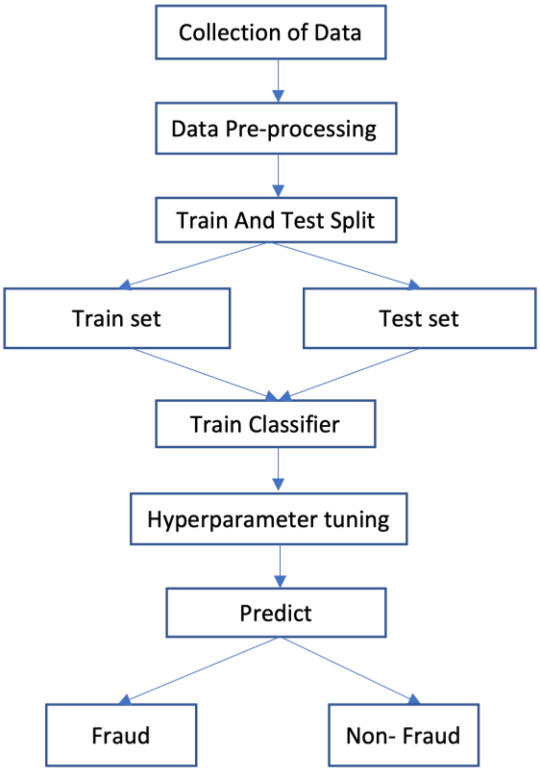

Figure 1: Proposed Model for Insurance Fraud Detection.

3.0.1 Data Overview

The dataset used in this study was obtained from Kaggle (*https* : *//www.kaggle.com/code/niteshyadav*3103*/insurance*− *fraud* − *detection* − *using* − 12 − *models/data*?*select* = *insurance claims.csv*).

The dataset consists of 1,000 auto incidents and auto insurance claims from Ohio, Illinois, and Indiana from January 1, 2015, to March 1, 2015. The dataset consists of 39 variables, each explained in Table 1 (the table content was not provided in the original document).

3.0.2 Data Pre-processing

Data pre-processing is one of the most impor- tant processes in Machine Learning. This is the transformation of data from its raw form into a format that machine learning models can un- derstand. Errors and missing values are almost inevitable in datasets, and managing these is- sues is the focus of this phase Peng (2020). For instance, outliers in the dataset were de- tected as depicted in Figure 2, and we cor- rectly categorized numeric and categorical data (Figure 3) before utilizing the Machine Learn- ing models to analyze our processed dataset. These steps are briefly illustrated below:

IV. MACHINE LEARNING ALGO- RITHMS

The different machine learning algorithms used in this paper are Random Forest (RF), De- cision Tree Classifiers (DTC), AdaBoost, K- Nearest Neighbor (KNN), and XGBoost (XG), which we will briefly review in detail below.

4.1 Decision Tree Classifiers

One of the most important features of DTCs is their capability to break down a complex decision-making process into a collection of simpler decisions, thus providing a solution that is often easier to interpret. Decision Tree Classifiers (DTCs) are successfully used in many diverse areas such as radar signal classi- fication, character recognition, remote sensing, medical diagnosis, expert systems, and speech recognition, to name only a few. The decision tree classifier is one approach to multistage de- cision making. The basic idea involved in any multistage approach is to break up a complex decision into a union of several simpler deci- sions, with the expectation that the final so- lution obtained this way would resemble the intended desired solution. This is illustrated in Figure **??**.

4.2 Random Forest

One of the most commonly used supervised from previous segments. More specifically, Random Forests are collections of decision trees that produce better forecast accuracy. It is essentially a group of decision trees, which is why it is termed a ”forest.” The basic idea is to create distinct decision trees based on in- dependent subsets of the dataset. The optimal split on n factors is determined at each node by randomly selecting one from the list of fea- tures. An illustration of the Random Forest algorithm is presented in Figure 5.

4.3 AdaBoost Algorithm

AdaBoost, also called Adaptive Boosting, is a technique in Machine Learning used as an En- semble Method. The most common algorithm used with AdaBoost is decision trees with one level, also known as Decision Stumps. This al- gorithm creates a model while assigning each data piece an equal weight. Then, it gives points that were incorrectly categorized larger weights. The next model gives more weight to all points with higher weights. If no lower error is received, it will continue to train the models. An illustration is shown in Figure 6.

4.4 K-Nearest Neighbor (KNN)

The KNN regressor is employed when there is no assumption about the frequency distribu- tion of the relationship between predictor vari- ables. This is achieved by averaging certain features in the same area. K-Nearest Neigh- bor has been found to be an effective classifi- cation technique, although this study doesn’t fully support that conclusion. The value of K is calculated using the square root of the total amount of data in the training dataset. In the KNN classification problem, the output would be a class to which the data model belongs, predicted by the majority vote of the k nearest neighbors. In the regression problem, the out- put would be a property estimation, typically a mean estimation of the k nearest neighbors. The Euclidean Distance can be used to deter- mine the nearest neighbor division:

$$ED = \sqrt{\sum_{i=1}^{N}(q_i - p_i)^2}, \qquad (1)$$

where ED is Euclidean distance, and p and q are points in n-space. The calculation’s ex-

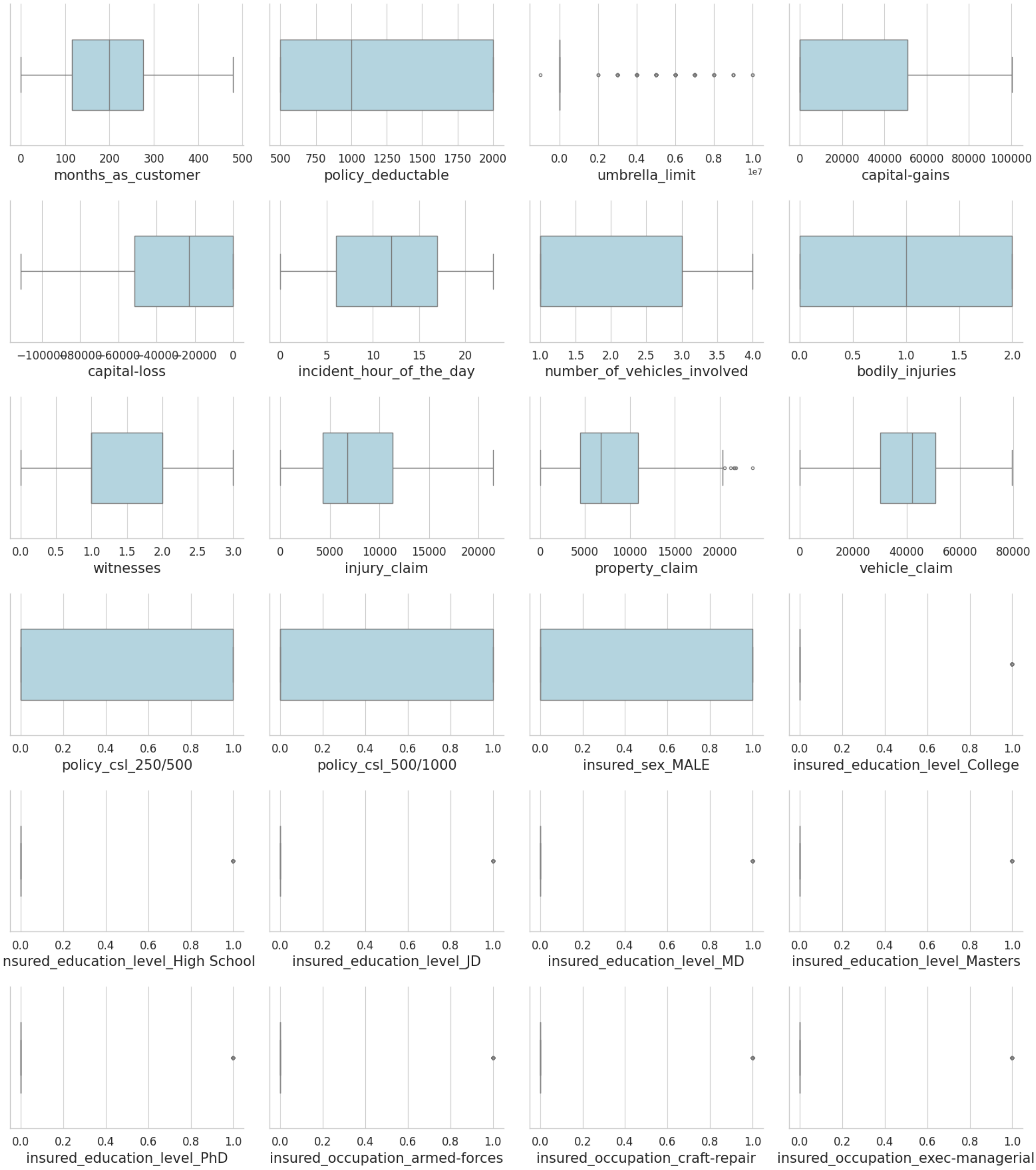

Figure 2: Visualizing outliers

pected accuracy depends heavily on the value of k. Smaller values of k will likely result in lower precision, especially in datasets with sig- nificant noise, since each instance in the train- ing set carries a larger weight throughout the decision process. The algorithm's performance decreases as k's value increases. Additionally, if the value is too high, the model may overfit, weakening the distinction between class bound- aries and causing reduced accuracy. As a general approach, it is advised to select k by ap- plying the formula:

$$k = \sqrt{n} \qquad (2)$$

The algorithm for K-Nearest Neighbor is illus- trated in Figure 7.

4.5 XGBoost

Extreme Gradient Boosting (XGBoost) is a distributed, scalable gradient-boosted decision

Figure 3: Encoding Categorical and Numerical data

tree (GBDT) machine learning framework. It is a top machine learning library for regres- sion, classification, and ranking problems, of- fering parallel tree boosting. Understanding XGBoost requires knowledge of machine learn- ing concepts and techniques, including super- vised learning, decision trees, ensemble learn- ing, and gradient boosting. In supervised ma- chine learning, a model is trained using algo- rithms to discover patterns in a dataset of features and labels, and the model is then used to predict labels for features in a new dataset. XGBoost is a scalable and extremely accurate gradient boosting solution that pushes the lim- its of computing power for boosted tree algo- rithms. It was developed primarily to enhance the performance and computational speed of machine learning models. Unlike GBDT's se- quential approach, XGBoost constructs trees in parallel. It employs a level-wise approach,

scanning over gradient values and assessing the quality of splits at each potential point in the training set using partial sums. The XGBoost algorithm is illustrated in Figure 8.

4.6 HYPERPARAMETER TUNING

Hyperparameter optimization in machine learning aims to find the hyperparameters of a given machine learning algorithm that deliver the best performance as measured on a valida- tion set. Unlike model parameters, hyperpa- rameters are set by the machine learning engi- neer before training. For example, the number of trees in a Random Forest is a hyperparam- eter, while the weights in a neural network are model parameters learned during training. Hy- perparameters can be considered as model set- tings to be tuned so that the model can opti- mally solve the machine learning problem. We use GridSearchCV for hyperparameter tuning.

V. EVALUATION METHOD

Evaluation methods are crucial for model com- parison and best model selection, as they assess the effectiveness of classifiers Hossin (2015). Since bias can sometimes be introduced for a majority class in classification problems due to imbalanced data, accuracy alone is not al- ways reliable Ganganwar (2012), Hanafy and Ming (2021). Car insurance claims exemplify imbalanced data because the majority of poli- cyholders don't engage in fraud. Therefore, if accuracy is the sole criterion, there will be bias against the fraud class. Consequently, several measurement techniques are employed, includ- ing F1-score, sensitivity, accuracy, and area un- der the curve (AUC). AUC might be a better metric when results need to balance sensitiv- ity and specificity, especially with imbalanced class distributions.

$$\mathrm{SENSITIVITY} = \frac{\mathrm{TP}}{(\mathrm{TP} + \mathrm{FN})} \qquad (3)$$

$$\mathrm{ACCURACY} = \frac{(\mathrm{TP} + \mathrm{TN})}{(\mathrm{TP} + \mathrm{FP} + \mathrm{TN} + \mathrm{FN})} \qquad (4)$$

$$\mathrm{SPECIFICITY} = \frac{\mathrm{TN}}{(\mathrm{FP} + \mathrm{TN})} \qquad (5)$$

$$\mathrm{PRECISION} = \frac{\mathrm{TP}}{(\mathrm{TP} + \mathrm{FP})} \qquad (6)$$

$$\text{F-MEASURE} = \frac{(2 \times \text{Precision} \times \text{Recall})}{(\text{Precision} + \text{Recall})} \qquad (7)$$

In these formulas, TP stands for the number of true positives, FP for false positives, TN for true negatives, and FN for false negatives. A higher Accuracy rating indicates better overall performance of the forecast. Accuracy mea- sures the percentage of predictions that are correct. Sensitivity pertains to the accuracy of fraud claims detection. Specificity refers to the ability to accurately identify legitimate claims. Precision measures the relevance of projected positives. The F1 score is the harmonic mean of precision and sensitivity. AUC is the over- all classifier performance metric Wu and Flach (2005) and is used to evaluate the model's over- all performance.

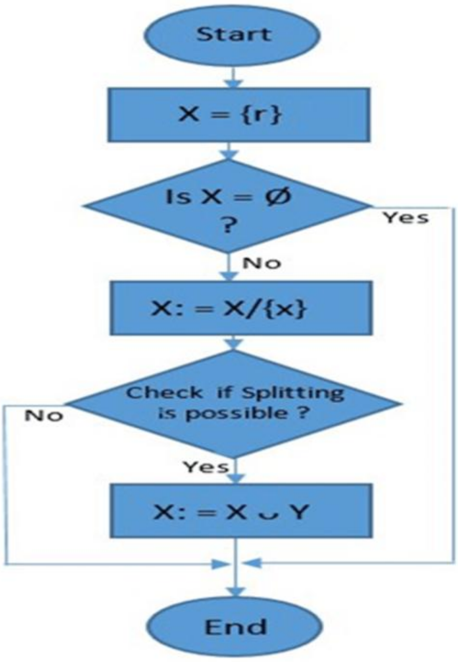

Figure 4: Flowchart of Decision Tree Classifiers

VI. RESULTS

Fraud must be addressed as it is a significant issue in today's society. To overcome these challenges, we can develop systems that detect fraud in provided data. These systems are created using various machine learning methods, including neural networks, naive Bayes, KNN, and random forests. In this paper, we've discussed different machine learning (ML) tech-

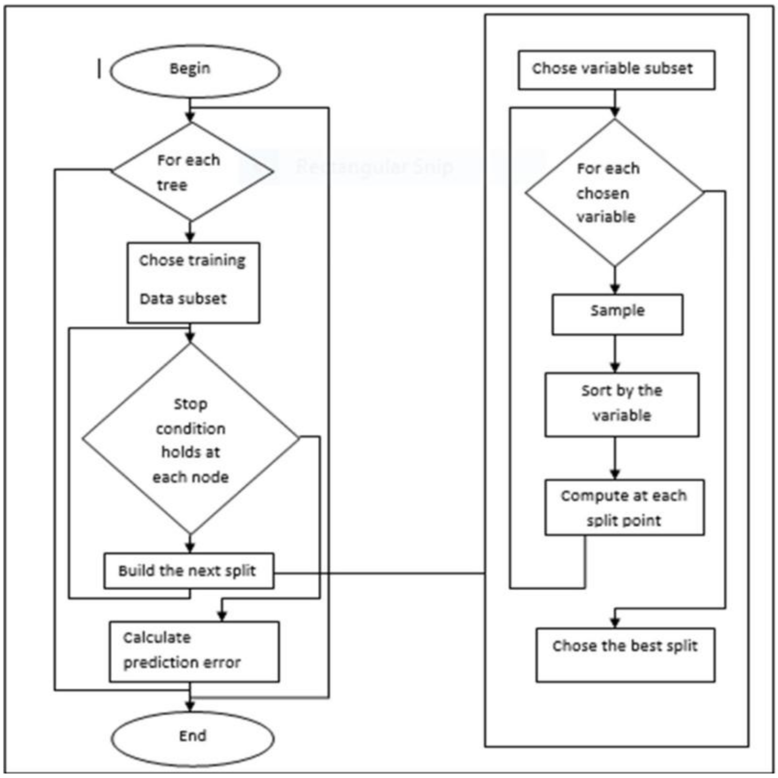

Figure 5: Flowchart of Random Forest Algorithm

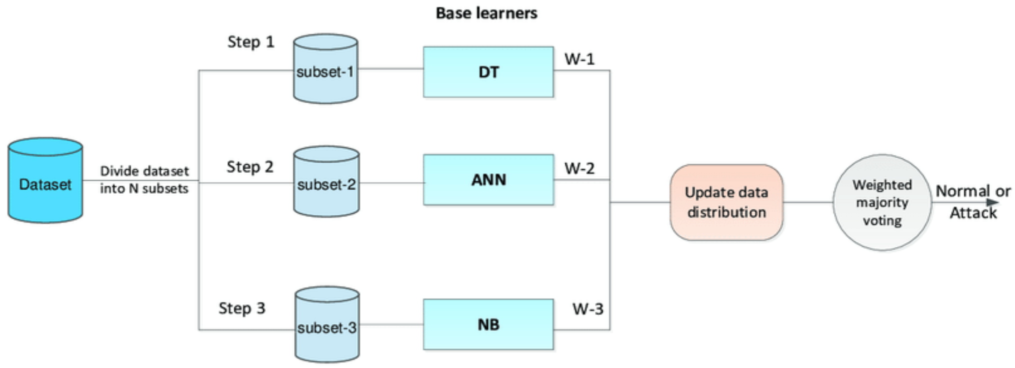

Figure 6: AdaBoost Flowchart

niques, their operation in systems, and their effectiveness in predicting fraud. These meth- ods are then compared using five criteria from various perspectives. To assess the perfor- mance of the machine learning algorithms on fraud discrimination, we randomly divide the data, using 80% for training and the remaining 20% for testing. We use the training data to train the algorithms, and the learned model is then used to forecast whether examples in the

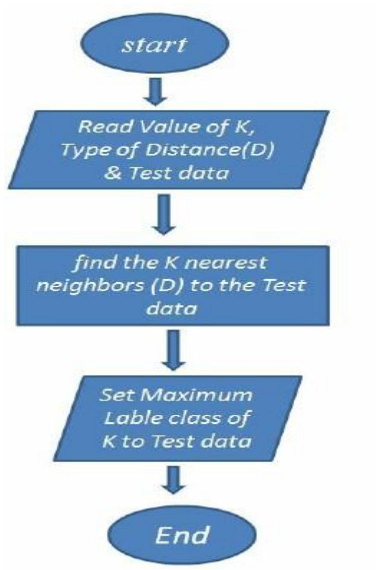

Figure 7: Flowchart of K-Nearest Neighbor Al- gorithm

test data are fraudulent. Six assessment tech- niques evaluate the performance of the mod- els on the testing data: accuracy, sensitivity, specificity, AUC, precision, and F1-score. Ad- ditionally, we optimize each machine learning model to ensure fair comparisons. It is neces- sary to point out that the study has a few draw- backs which include limited sample size. This could reduce the performance of the model since larger datasets increase the stability of statistical models and generalize better by in- cluding a larger fraction of the actual population.

Table 1: Comprehensive Performance Compar- ison of Models

Model	Train Acc	Test Acc	N Prec
K-Nearest Neighbors	0.755	0.748	0.75
Decision Tree	0.816	0.696	0.85
Random Forest	0.981	0.784	0.81
AdaBoost	0.816	0.796	0.87
XGBoost	0.932	0.712	0.82
Voting Classifier	0.816	0.796	0.87

VII. Discussion

Our analysis of multiple machine learning al- gorithms for insurance fraud detection reveals several important patterns with significant im- plications for practical applications. Figure 10 shows that the dataset exhibits a notable class imbalance (187 non-fraud vs. 63 fraud cases in the test set), reflecting real-world conditions where fraudulent claims represent the minority. This imbalance substantially impacts model performance, as demonstrated by KNN's (Fig- ure 10 top left) complete failure to detect fraud (0% Y-class recall) despite achieving a re- spectable overall accuracy of 74.8%. This find- ing exemplifies the "majority class bias" prob- lem, where algorithms optimize overall accu- racy by simply predicting the dominant class, and highlights why accuracy alone is an inad- equate metric for fraud detection systems.

Across all models, we observe a clear trade- off between precision and recall. Random For- est achieves high precision for fraud detec- tion (0.64) but low recall (0.33), meaning it rarely misclassifies legitimate claims as fraud- ulent but fails to identify many actual fraud cases. Conversely, Decision Tree shows lower precision (0.43) but higher recall (0.62), cap- turing more fraud cases at the expense of more false positives. Furthermore, Figure 9 shows that AdaBoost and Voting Classifier offer the most balanced trade-off (0.59 precision, 0.62 recall), making them potentially more suit- able for real-world deployment. This precision- recall trade-off has direct business implica- tions: high precision reduces false accusations of fraud (which can damage customer rela- tionships), while high recall ensures more ac- tual fraud cases are caught (reducing financial losses).

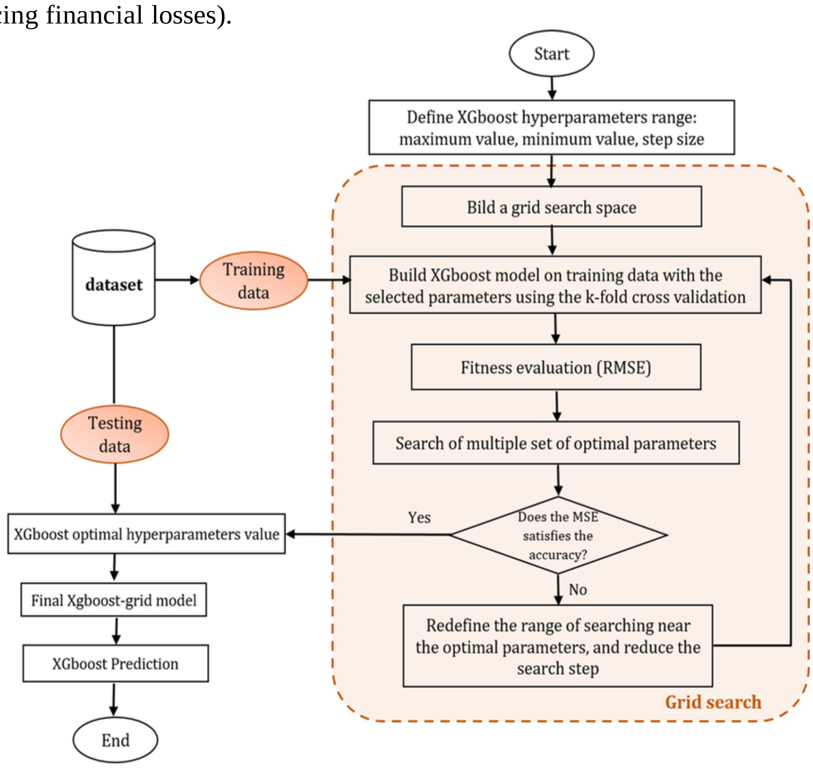

Figure 8: Flowchart of XGBoost.

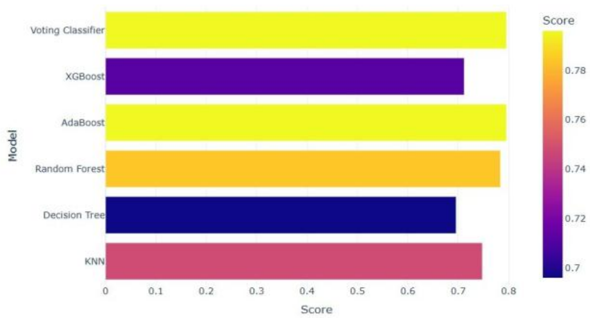

Figure 9: Comparison of model scores

The relationship between model complexity and performance is evident in our results. Sim- ple models like KNN struggle with nuanced classification tasks in imbalanced data, while highly flexible models such as Random For- est and XGBoost show signs of overfitting, as indicated by the substantial gap between their training and testing accuracy (0.981 vs. 0.784 for Random Forest; 0.932 vs. 0.712 for XGBoost). This suggests these models are memorizing training data rather than learning generalizable patterns, and that regularization techniques might improve their performance on new data. Ensemble methods (AdaBoost, Voting Classifier) provide more robust perfor- mance by combining multiple models, result- ing in the highest overall test accuracy (79.6%) and the best F1-score for detecting fraudulent claims (0.60).

The correlation heatmap reveals that claim- related variables (total claim amount, in- jury claim, property claim, vehicle claim) are highly correlated. This suggests that fraud- sters may simultaneously inflate claims across multiple categories, a pattern that could be exploited for improved detection. Addition-

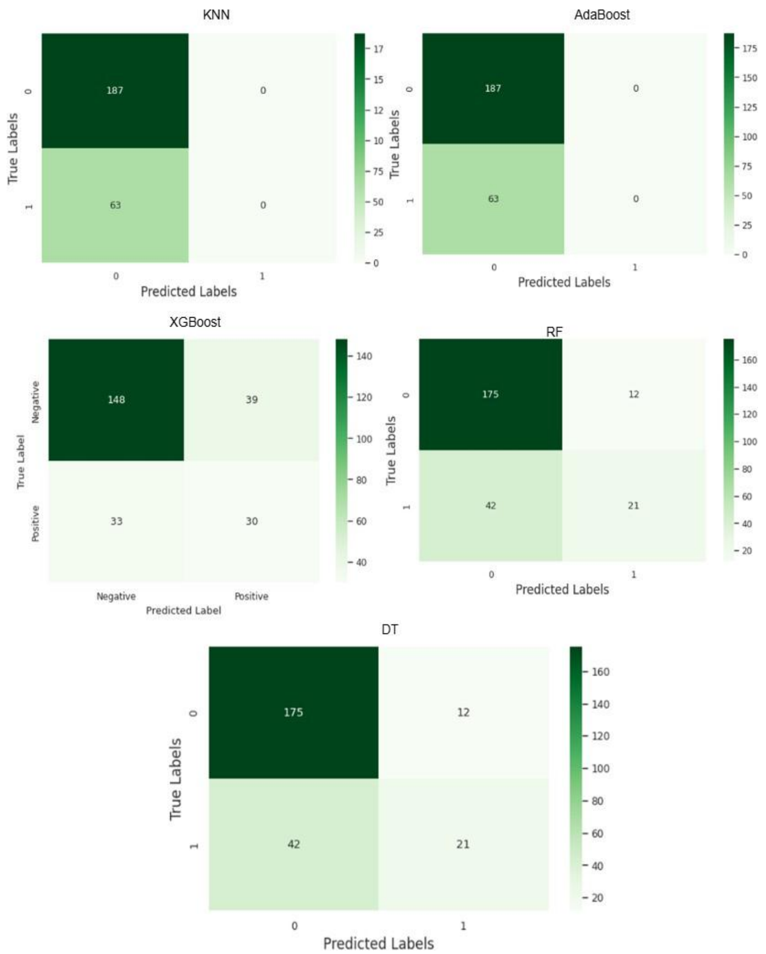

Figure 10: Confusion matrices of various models

ally, the strong correlation (0.92) between age and months as customer indicates that older customers tend to have longer relationships with the insurance company, potentially mak- ing them less likely to commit fraud.

VIII. Conclusion

This research demonstrates that effective in- surance fraud detection requires models that balance overall accuracy with minority class detection capabilities. Our findings indicate that AdaBoost emerges as the optimal solution for insurance fraud detection, providing the best balance between detecting legitimate and fraudulent claims with a test accuracy of 79.6% and fraud class F1-score of 0.60. Model selection in this domain should prioritize class- specific metrics over aggregate accuracy, par- ticularly given the significant class imbalance inherent in insurance fraud data.

Ensemble methods consistently outperform individual classifiers in our study, likely due to their ability to combine diverse decision bound- aries and reduce variance. The correlation be- tween claim-related variables offers a potential avenue for feature engineering in future fraud detection systems. Additionally, our results highlight that proper hyperparameter tuning is essential to mitigate overfitting, particularly for complex models like Random Forest and XGBoost.

Our study has several limitations that sug- gest directions for future research. The lim- ited sample size (1,000 records) may not cap- ture the full diversity of fraud patterns. Fu- ture work should incorporate larger datasets spanning longer periods and more geographic regions. More sophisticated handling of class imbalance through techniques like SMOTE (Synthetic Minority Oversampling Technique) could potentially improve model performance. Deep learning approaches were not explored in this study but could offer improved per- formance, particularly for identifying complex fraud patterns. Time-based validation would better simulate real-world conditions where models must detect fraud in future periods based on historical data. Finally, cost-sensitive evaluation that incorporates the financial im- pact of false positives versus false negatives would provide a more business-relevant assess- ment of model performance.

These insights can guide insurance compa- nies in implementing more effective fraud de- tection systems that better balance the dual objectives of minimizing false accusations while maximizing fraud detection rates. By address- ing the limitations identified in our study, fu- ture research can build upon our findings to develop even more effective insurance fraud detection systems that help reduce industry losses while treating customers fairly.

References

[1]. Awoyemi, J. O., Adetunmbi, A. O., & Oluwadare, S. A. (2017). Credit card fraud detection using machine learning techniques: A comparative analysis. In 2017 international conference on com- puting networking and informatics (ic- cni) (pp. 1–9).

[2]. Carcillo, F., Le Borgne, Y.-A., Caelen, O., Kessaci, Y., Oblé, F., & Bontempi, G. (2021). Combining unsupervised and su- pervised learning in credit card fraud de- tection. Information sciences, 557 , 317– 331.

[3]. Dal Pozzolo, A., Caelen, O., Le Borgne, Y.-A., Waterschoot, S., & Bontempi, G. (2014).

[4]. Learned lessons in credit card fraud de- tection from a practitioner perspective. Expert systems with applications, 41 (10), 4915–4928.

[5]. Eshghi, A., & Kargari, M. (2019). Introduc- ing a new method for the fusion of fraud evidence in banking transactions with re- gards to uncertainty. Expert Systems with Applications, 121 , 382–392.

[6]. French Agency for the Fight against Insur- ance Fraud, A. (2021). Fight against fraud conference. ALFA Homepage, https://www.alfa.asso.fr/.

[7]. Ganganwar, V. (2012). An overview of classification algorithms for imbal- anced datasets. International Journal of Emerging Technology and Advanced En- gineering, 2 (4), 42–47.

[8]. Gottschalk, P. (2010). Categories of financial crime. Journal of financial crime.

[9]. Hanafy, M., & Ming, R. (2021). Machine learning approaches for auto insurance big data. Risks, 9 (2), 42.

[10]. Hossin, M. S., Mohammad. (2015). A review on evaluation metrics for data classifica- tion evaluations. International Journal of Data Mining and Knowledge Manage- ment Process 5: 1 .

[11]. Insurance Fraud Bureau of Australia, I. (2021). Insurance fraud. Insurance Fraud Bureau of Australia Homepage, https://ifba.org.au.

[12]. Itri, B., Mohamed, Y., Mohammed, Q., & Omar, B. (2019). Performance com- parative study of machine learning al- gorithms for automobile insurance fraud detection. In 2019 third international conference on intelligent computing in data sciences (icds) (pp. 1–4).

[13]. Kim, Y. J., Baik, B., & Cho, S. (2016). Detect- ing financial misstatements with fraud intention using multi-class cost- sensitive learning. Expert systems with applica- tions, 62 , 32–43.

[14]. Ngai, E. W., Hu, Y., Wong, Y. H., Chen, Y., & Sun, X. (2011). The application of data mining techniques in financial fraud de- tection: A classification framework and an academic review of literature. Deci- sion support systems, 50 (3), 559–569.

[15]. Peng, . N., Y. (2020). An empirical overview of nonlinearity and overfitting in machine learning using covid-19 data. Chaos, Soli- tons and Fractals, Article 110055.

[16]. Raghavan, P., & El Gayar, N. (2019). Fraud detection using machine learning and deep learning. In 2019 international con- ference on computational intelligence and knowledge economy (iccike) (pp. 334– 339).

[17]. Sheshasaayee, A., & Thomas, S. S. (2018). Usage of r programming in data analyt- ics with implications on insurance fraud detection. In International conference on intelligent data communication tech- nologies and internet of things (pp. 416–421).

[18]. S.Jordon. (2016). Insurance fraud: 'its all over the place' and you should care about it officials say. Omaha World-Herald , https://omaha.com/business/.

[19]. Varmedja, D., Karanovic, M., Sladojevic, S., Arsenovic, M., & Anderla, A. (2019). Credit card fraud detection-machine learning methods. In 2019 18th inter- national symposium infoteh-jahorina (in- foteh) (pp. 1–5).

[20]. Waghade, S. S., & Karandikar, A. M. (2018). A comprehensive study of healthcare fraud detection based on machine learning. In- ternational Journal of Applied Engineer- ing Research, 13 (6), 4175–4178.

[21]. Wang, Y., & Xu, W. (2018). Leveraging deep learning with lda-based text analytics to detect automobile insurance fraud. Deci- sion Support Systems, 105 , 87–95.

[22]. Wu, S., & Flach, P. (2005). A scored auc metric for classifier evaluation and selection. In Second workshop on roc analysis in ml, bonn, germany.

[23]. Xu, W., Wang, S., Zhang, D., & Yang, B. (2011). Random rough subspace based neural network ensemble for insurance fraud detection. In 2011 fourth in- ternational joint conference on compu- tational sciences and optimization (pp. 1276–1280).