AI Integration in Decision-Support Systems for Insurance Fraud Detection

Thang Sian Kop

Abstract

This paper extends a prior proposal to incorporate artificial intelligence (AI) into the operations of an insurance company by developing and analyzing a decision-support framework for fraud detection. Using a simulated dataset representative of real insurance claims, the study applies three machine learning models, Decision Tree, Random Forest, and Multilayer Perceptron (MLP), to evaluate the effectiveness of predictive analytics in identifying fraudulent activity. Evaluation metrics such as ROC-AUC and Precision-Recall (PR) curves are used to assess model performance and reliability. The study highlights interpretability, calibration, and human oversight as critical factors in adopting AI for operational decision-support systems. Finally, the discussion explores ethical and governance considerations to ensure responsible and trustworthy deployment.

AI Integration in Decision-Support Systems for Insurance Fraud Detection

1. Introduction

Insurance fraud poses a substantial financial burden on insurers and policyholders worldwide. Traditional rule-based systems often struggle to adapt to new patterns of fraudulent behavior, leading to inefficiencies and losses. With advances in AI and machine learning, insurers now can automate fraud detection processes while preserving human oversight and accountability.

The current research extends that proposal by developing a technical, data-driven framework that integrates AI into the company's decision-support system. The study focuses on building interpretable models that balance automation with transparency, enabling analysts to make better, faster, and more data-informed decisions.

The objectives of this research are threefold:

- To evaluate the performance of common AI models on an imbalanced fraud dataset.
- To analyze trade-offs between accuracy, interpretability, and operational scalability.
- To demonstrate how predictive models can be embedded into a decision-support workflow that assists but does not replace human experts.

2. Literature Review

Artificial intelligence in the insurance sector has rapidly evolved from experimental pilots to core business functions. PwC (2023) identified that insurers leveraging predictive analytics can reduce claim-processing time by 40% and improve fraud detection accuracy by more than 25%. Similarly, the Cloud Security Alliance (2023) emphasized the importance of secure data governance when deploying AI systems, warning that breaches in model management can lead to reputational and financial damage.

Scholarly research aligns with these industrial insights. Doshi-Velez and Kim (2017) advocate for interpretable machine learning as a prerequisite for responsible deployment, arguing that decision-support systems must remain explainable to stakeholders. Ribeiro, Singh, and Guestrin (2016) introduced methods such as LIME to explain individual model predictions, further reinforcing the need for transparency. In operational contexts, Pedregosa et al. (2011) demonstrated the use of scikit-learn pipelines to ensure reproducibility and consistency in applied machine learning workflows.

Organizationally, Ashkenas and Matta (2021) describe a framework for scaling successful pilot projects, which can be adapted to AI adoption in insurance. Their approach emphasizes small-scale experimentation followed by progressive rollout, enabling companies to validate AI benefits before full integration. Sjödin et al. (2021) add that AI-driven innovation depends not only on algorithms but also on organizational culture, data maturity, and employee training.

Taken together, these studies support a hybrid model of AI adoption, one that integrates technical innovation with human oversight, transparency, and ethical governance.

3. Methods

3.1 Data Simulation

A synthetic dataset of 25,000 insurance claims was generated to model real-world operational data. Each record includes features such as policy age, claim amount, prior claims, provider tenure, claim type, and region. A probabilistic mechanism simulates fraud likelihood, ensuring an approximately 8% fraud rate consistent with industry statistics. This approach maintains privacy while allowing for realistic experimentation.

3.2 Preprocessing

Data preprocessing was performed using a unified ColumnTransformer to standardize numerical variables and one-hot encode categorical features. This ensures that all models receive consistent and comparable inputs, facilitating fair evaluation and reproducibility. Feature scaling also mitigates the impact of extreme values in claim amounts.

3.3 Model Selection

Three supervised learning algorithms were implemented:

- Decision Tree (CART): Provides high interpretability and visual transparency, allowing investigators to trace the logic behind predictions.
- Random Forest: Offers improved generalization through ensemble averaging, reducing overfitting while maintaining feature importance insights.
- Multilayer Perceptron (MLP): A neural network architecture capable of modeling non-linear relationships within data.

All models were trained on 80% of the dataset, with 20% reserved for testing. Stratified sampling ensured that the minority (fraudulent) class was proportionally represented.

3.4 Evaluation Metrics

Performance was assessed using the Receiver Operating Characteristic Area Under Curve (ROC-AUC) and the Precision-Recall Area Under Curve (PR-AUC). ROC-AUC captures the model's overall discriminatory ability, while PR-AUC provides greater insight into rare-event classification such as fraud detection. Confusion matrices were also analyzed to understand the balance between false positives (legitimate claims flagged) and false negatives (fraudulent claims missed).

4. Results

Among the models tested, the Random Forest achieved the highest ROC-AUC and PR-AUC scores, confirming the robustness of ensemble methods for complex datasets. The Decision Tree model, while less accurate, proved highly interpretable, making it useful in early-stage audits or compliance reporting. The MLP model delivered strong performance but required careful tuning and computational resources.

A comparative summary is as follows:

- Random Forest: ROC-AUC ≈ 0.93, PR-AUC ≈ 0.74
- Decision Tree: ROC-AUC ≈ 0.86, PR-AUC ≈ 0.61
- MLP: ROC-AUC ≈ 0.91, PR-AUC ≈ 0.71

Threshold tuning further demonstrated how operational policies can prioritize either precision or recall depending on risk appetite. For instance, a higher threshold favors accuracy (reducing false positives), while a lower threshold improves coverage of fraudulent claims. Such flexibility allows integration into human-in-the-loop decision-making workflows.

5. Discussion

The findings validate the feasibility of AI-driven fraud detection as part of a decision-support system rather than a fully autonomous solution. Random Forest models, while less transparent, can be augmented with explainability tools such as SHAP or LIME to maintain accountability. Decision Trees serve as a bridge between technical models and non-technical stakeholders by providing easily interpretable visualizations.

From an operational perspective, integrating AI into existing workflows requires alignment between data science, IT infrastructure, and claims departments. Analysts should be trained to interpret AI outputs responsibly, understanding that predictive scores indicate probability, not certainty, of fraud.

Furthermore, the study underscores the importance of continuous monitoring. AI models must be retrained periodically to capture changes in fraud patterns, regulatory environments, or consumer behavior. As Doshi-Velez and Kim (2017) emphasized, interpretability and auditability are essential to sustaining trust in AI systems.

6. Ethical and Governance Considerations

Ethical AI in insurance must balance automation with fairness, privacy, and accountability. The Cloud Security Alliance (2023) outlines frameworks for risk management, advocating data encryption, access control, and regular audits of AI systems. Biased training data can inadvertently target specific demographics, leading to unfair claim denials. Therefore, fairness audits and transparency documentation should accompany every system deployment.

Explainability tools, such as feature importance visualization, help stakeholders understand why a claim was flagged as suspicious. This transparency is critical not only for customer trust but also for regulatory compliance. Additionally, human review must remain a core component of the workflow to provide appeal and recourse for claimants affected by automated decisions.

7. Conclusion and Future Work

This research demonstrates a scalable, interpretable, and ethical approach to integrating AI into insurance decision-support systems. Random Forest models provide high accuracy, while Decision Trees enhance interpretability. Together, they form a foundation for a hybrid decision-support framework where AI assists rather than replaces human judgment. By embedding AI models within an analyst-driven workflow, organizations can improve efficiency and fraud detection rates while maintaining the transparency and fairness required in regulated industries.

Future research should focus on incorporating cost-sensitive learning techniques that account for the financial implications of false positives and false negatives. Extending the dataset with temporal or sequential claim histories could also enable the application of time-series models such as LSTMs or Transformers, providing dynamic risk assessment over time. Furthermore, integrating explainability tools like SHAP or LIME into operational dashboards would empower analysts to understand model predictions at a granular level, ensuring decisions remain interpretable and accountable.

From an implementation perspective, developing cloud-based decision-support interfaces using secure APIs and model monitoring frameworks (e.g., MLflow or EvidentlyAI) would ensure traceability, scalability, and compliance with data governance standards. Regular model audits, fairness assessments, and stakeholder training should accompany any deployment to sustain trust and mitigate risk.

Ultimately, the study underscores that AI is most powerful when used not as a replacement for human expertise but as an augmentation of it. When implemented responsibly, AI-driven decision-support systems can help insurers detect fraud more accurately, allocate resources more effectively, and deliver fairer outcomes for customers.

References

Ashkenas, R., & Matta, N. (2021, January 8). How to scale a successful pilot project. Harvard Business Review. https://hbr.org/2021/01/how-to-scale-a-successful-pilot-project

Cloud Security Alliance. (2023, August 11). AI security and risk management. https://cloudsecurityalliance.org/blog/2023/08/11/ai-security-and-risk-management/

Doshi-Velez, F., & Kim, B. (2017). Towards a rigorous science of interpretable machine learning. arXiv. https://arxiv.org/abs/1702.08608

Kop, T. S. (2023). Incorporating Artificial Intelligence into CSU Insurance Company (Course proposal).

Pedregosa, F., et al. (2011). Scikit-learn: Machine learning in Python. Journal of Machine Learning Research, 12, 2825–2830. http://jmlr.org/papers/v12/pedregosa11a.html

PwC. (2023). Auto insurance AI analytics. https://www.pwc.com/us/en/library/case-studies/auto-insurance-ai-analytics.html

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). “Why should I trust you?”: Explaining the predictions of any classifier. In Proceedings of KDD (pp. 1135–1144). https://doi.org/10.1145/2939672.2939778

Sjödin, D., Parida, V., Leksell, M., & Petrovic, A. (2021). Data-driven business model innovation—AI's role. Journal of Business Research, 134, 574–587. https://doi.org/10.1016/j.jbusres.2021.05.021