sciendo

# Harnessing Artificial Intelligence for Risk Assessment and Fraud Detection in Insurance: A Modern Approach to Predictive Modelling

**Mihai VRISCU**
*Bucharest Academy of Economic Studies, Bucharest, Romania*
*mvriscu@yahoo.com*

**Abstract.** *This research investigates the potential of artificial intelligence (AI) to enhance two of the most critical processes in the insurance industry: risk assessment and fraud detection. As insurers face increasing pressure to process complex data, minimize losses, and provide more personalized services, traditional statistical models such as logistic regression often prove insufficient in capturing the nonlinear relationships and subtle patterns embedded in client behavior and claim history. The central research question addressed in this study is: How can artificial intelligence improve risk assessment and fraud detection in insurance compared to traditional statistical approaches? To answer this question, a simulation-based experimental methodology was employed. A synthetic dataset containing 10,000 insurance client profiles was generated, incorporating realistic demographic, behavioral, and policy-related variables. Several machine learning models were implemented, including supervised algorithms (Logistic Regression, Random Forest, XGBoost, Neural Networks, SVM) and unsupervised methods (Isolation Forest, Autoencoder, K-Means), with performance evaluated using accuracy, F1 score, precision, and recall. Model training and validation were conducted using cross-validation techniques and hyperparameter optimization to ensure robustness and generalizability. The results demonstrate a clear performance advantage of AI-based models over traditional statistical methods. XGBoost emerged as the best-performing model, achieving the highest accuracy (87.3%), F1 score (0.86), and strong precision and recall, confirming its robustness in both classification and detection tasks. Random Forest and Neural Networks also performed well, while Logistic Regression lagged significantly behind. Unsupervised models such as Isolation Forest and Autoencoders proved particularly useful in fraud detection, offering high recall rates suitable for anomaly detection and early-stage screening. The integration of multiple models into a hybrid AI architecture is recommended to balance precision with sensitivity. In conclusion, the findings of this research affirm that artificial intelligence provides a powerful and scalable solution for modern insurance analytics. AI not only improves the precision and efficiency of risk assessment and fraud detection but also enables dynamic, data-driven decision-making that traditional models cannot replicate. As the insurance industry becomes increasingly digitized, AI adoption represents not merely an opportunity but a strategic necessity for organizations seeking to optimize operations and maintain a competitive edge.*

**Keywords**: Artificial Intelligence; Insurance Risk Assessment; Predictive Modeling; Fraud Detection; Risk Classification; Insurance Analytics.

## Introduction

In an increasingly complex and unpredictable world, the concept of risk lies at the very heart of the insurance industry. Insurance, by definition, exists as a financial mechanism to transfer and mitigate risk, offering individuals and institutions a means to protect themselves against uncertain and potentially catastrophic events. Traditionally, this process has relied on actuarial science and statistical modeling, drawing on historical data and probability theory to evaluate exposure and set premiums. While these classical approaches have formed the backbone of insurance decision-making for decades, they are increasingly being challenged by the volume, velocity, and variety of

DOI: 10.2478/picbe-2025-0179

modern data, as well as by the growing complexity of consumer behavior, environmental threats, and economic volatility.

At the same time, the field of artificial intelligence (AI) has emerged as a disruptive and transformative force across numerous industries, including healthcare, finance, and logistics. Its ability to process vast datasets, identify complex patterns, and generate predictive insights in real-time has opened new avenues for decision-making that go far beyond human or rule-based capabilities. In the insurance sector, AI promises to revolutionize not only how risk is assessed, but also how fraud is detected, how policies are priced, and how claims are processed. By integrating machine learning algorithms and advanced data analytics into their operations, insurers can transition from reactive to proactive strategies—anticipating risk rather than merely responding to it.



The intersection between AI and insurance risk management presents a fertile ground for academic and practical exploration. Risk, by nature, is probabilistic and multifaceted, influenced by a range of demographic, behavioral, economic, and environmental factors. Artificial intelligence, with its capacity to model non-linear relationships and continuously learn from new data, is uniquely positioned to enhance risk classification and portfolio optimization. Moreover, the detection of fraud—a persistent challenge for insurers worldwide—can benefit greatly from AI's anomaly detection and pattern recognition capabilities, enabling faster and more accurate identification of fraudulent claims, even when they deviate from known patterns.

Despite its promise, the integration of AI into insurance analytics is not without challenges. Issues of data quality, algorithmic transparency, regulatory compliance, and ethical fairness are central to any discussion on AI deployment in financial services. Furthermore, the shift toward AI-driven systems requires a rethinking of traditional actuarial roles, business processes, and technological infrastructure.

This study aims to contribute to this growing field by examining how artificial intelligence can improve the core tasks of risk assessment and fraud detection in the insurance sector. Through a comparative analysis of AI-based models and traditional methods, this research seeks to evaluate the performance, reliability, and operational implications of adopting AI tools in a simulated insurance environment. By doing so, it provides a foundation for understanding the current capabilities and future potential of AI in reshaping risk management within the insurance industry.

## Literature review

### *The Evolution of Risk Assessment in Insurance: From Traditional Models to AI Integration*

The field of insurance has always been inherently tied to the concept of risk. Historically, risk assessment relied heavily on statistical models rooted in actuarial science, which utilized predefined variables and probability distributions to estimate potential losses. These traditional methodologies, while foundational, were inherently limited by their reliance on assumptions of normality, linearity, and stationarity, which often failed to capture the complexity and dynamism of real-world insurance environments (Canhoto and Clear, 2020). In recent decades, however, the exponential growth in data availability and computational power has triggered a paradigm shift in how risk is evaluated, enabling the transition from static, rule-based systems to more adaptive, data-driven approaches.

Artificial intelligence (AI), particularly through its subfields of machine learning and deep learning, has emerged as a transformative force in the insurance industry (Logunova, 2024). Unlike conventional models, AI systems are capable of learning from vast datasets, identifying intricate patterns, and making probabilistic inferences without being explicitly programmed for specific

rules. This capability is particularly crucial in a sector where risk variables are becoming increasingly multi-dimensional—ranging from climate change impacts to behavioral data captured through IoT devices and mobile applications (Reis et al., 2020). The incorporation of AI has enabled insurers to move from reactive underwriting based on historical averages to proactive, individualized risk predictions that adjust in real time (Paruchuri, 2021).



Moreover, predictive analytics powered by AI enhances insurers' ability to stratify risk among heterogeneous policyholders. Advanced algorithms can now incorporate variables that were previously considered non-quantifiable, such as social media behavior, geolocation data, or sentiment extracted from text (Lee and Shin, 2019). These capabilities allow for the dynamic modeling of client profiles, resulting in more accurate pricing, improved portfolio management, and minimized exposure to high-risk segments. As such, AI is not merely an incremental improvement to risk assessment—it represents a foundational redefinition of the discipline, reshaping the actuarial role and the entire underwriting ecosystem (Miller and Fang, 2023).

***Artificial Intelligence in Fraud Detection: Techniques, Challenges, and Emerging Practices***Fraud detection in the insurance sector presents a uniquely challenging problem characterized by high dimensionality, class imbalance, and the presence of adversarial behavior (Kibria and Sevkli, 2021). Traditionally, insurers have relied on manual audits and rule-based expert systems to identify suspicious claims. While these systems are useful for detecting known patterns of fraud, they struggle to adapt to new, evolving schemes that fraudsters employ. In contrast, AI-based models, particularly those employing supervised and unsupervised learning techniques, have demonstrated significant efficacy in identifying anomalous behavior and uncovering hidden relationships that may indicate fraudulent activity (Đorđe Krivokapić and Nikolić, 2022; Busu, 2018).

Machine learning algorithms such as decision trees, support vector machines, and random forests are now widely used for supervised fraud detection tasks. These models learn from labeled datasets to differentiate between legitimate and suspicious claims, constantly improving as new data is ingested (Cekuls, 2023). On the other hand, unsupervised models—such as clustering, autoencoders, or isolation forests—are invaluable for detecting unknown or rare fraud types, especially in scenarios where labeled fraudulent data is scarce (Gatzert and Kolb, 2013). Additionally, recent advancements in natural language processing (NLP) have allowed insurers to analyze unstructured textual data, such as claims descriptions or customer emails, for inconsistencies and deceitful narratives (Kietzmann and Pitt, 2019).

Despite these advancements, several challenges persist. One of the primary concerns is the issue of false positives, which can lead to unnecessary claim delays and customer dissatisfaction. Moreover, the dynamic nature of fraud—where perpetrators continuously evolve their strategies in response to detection methods—requires AI systems that are not only accurate but also agile and continuously updated (Tamara Adel Al-Maaitah et al., 2024). Ethical concerns also arise in the context of AI-driven fraud detection, particularly with regard to transparency, fairness, and the potential for algorithmic bias. For instance, if historical data contains biases—such as disproportionate scrutiny of certain demographic groups—AI models may inadvertently perpetuate these injustices (Yermish et al., 2020).

Nevertheless, the application of AI in fraud detection is rapidly maturing. Leading insurance companies are adopting hybrid approaches that combine human expertise with automated decision systems, creating feedback loops where analysts validate AI-generated alerts, which in turn refine the models further. The integration of blockchain and federated learning

techniques is also being explored to enhance data privacy while maintaining model performance across decentralized datasets (O'Neill, 2019). Overall, AI represents a powerful tool in the ongoing battle against insurance fraud, offering both scalability and precision that far surpass traditional systems.



## Methodology

The methodological foundation of this study is grounded in a quantitative and simulation-based research approach aimed at assessing the applicability and effectiveness of artificial intelligence in improving risk assessment and fraud detection in the insurance industry. Given the sensitive nature of insurance data and the limited public availability of comprehensive datasets that include both risk-related variables and fraud labels, the decision was made to employ synthetically generated data that mimics real-world characteristics. This allowed for the implementation and evaluation of various machine learning algorithms in a controlled environment, while still preserving the variability, dimensionality, and complexity inherent in real insurance scenarios.

The primary research question guiding this investigation is as follows: *How can artificial intelligence improve risk assessment and fraud detection in insurance compared to traditional statistical approaches?* This question serves as a basis for exploring whether AI-powered predictive modeling can outperform conventional actuarial techniques, not only in terms of accuracy and efficiency, but also in adaptability to changing client behaviors and fraud schemes.

The focus is placed on two core insurance challenges—risk classification for underwriting purposes and the early identification of fraudulent claims—both of which directly impact the operational sustainability and profitability of insurance providers.

To simulate the insurance environment, a dataset comprising 10,000 synthetic client profiles was generated using Python programming tools such as scikit-learn and the Faker library. Each profile was constructed to include a variety of demographic, behavioral, and policy-related features, such as age, income, vehicle type, claim history, policy duration, and geographic risk exposure. Moreover, behavioral risk indicators—such as a telematics-based driving score—were introduced to reflect modern data sources increasingly used in insurance analytics. For the fraud detection component, a binary label was assigned to indicate whether a claim was fraudulent, simulating a scenario where 5% of the claims exhibit suspicious characteristics. This class imbalance reflects the typical distribution observed in real-world insurance datasets, where fraudulent cases are rare but significantly impactful.

A suite of machine learning models was deployed to process and analyze the data, encompassing both supervised and unsupervised learning paradigms. For the risk assessment task, models such as logistic regression, random forests, gradient boosting (via XGBoost), and neural networks (using multi-layer perceptrons) were trained to classify clients into different risk categories and predict expected premium pricing. The models were evaluated based on classification accuracy, F1-score, precision, and the area under the ROC curve, offering a robust comparison of performance across multiple metrics. For fraud detection, both supervised models (such as support vector machines) and unsupervised models (including isolation forests, autoencoders, and k-means clustering) were applied, with a focus on anomaly detection and pattern recognition. Given the class imbalance inherent to fraud detection tasks, emphasis was placed on metrics such as precision-recall trade-offs and the Matthews correlation coefficient, rather than simple accuracy.

The technical implementation of the models was carried out in Python using advanced data science libraries such as pandas, scikit-learn, keras, and matplotlib. Model training and evaluation were performed using Jupyter Notebooks in a Google Colab environment, which allowed access to cloud-based GPUs for efficient deep learning computations. To ensure model reliability and to mitigate overfitting, hyperparameters were tuned using grid search, and five-fold cross-validation was implemented throughout the experimental process. This methodological rigor ensured that the models generalized well across unseen data and that the findings could offer practical insights relevant to real-world insurance applications.



While the methodology provides a comprehensive framework for assessing the role of AI in insurance analytics, it is not without limitations. The use of synthetic data, while realistic in structure, may not fully capture the intricacies, noise, and adversarial tactics present in actual fraud cases. Moreover, the assumption of clean, structured, and fully available data may not hold in operational contexts where missing values, reporting delays, or inconsistent formats are common. Nonetheless, the methodology provides a controlled setting to isolate the performance of AI algorithms and to generate comparative insights that can inform future research and industry practices, particularly once real-world datasets become accessible.

## Results and discussions

The application of artificial intelligence models on the simulated insurance dataset has yielded a rich array of insights into the comparative effectiveness of different machine learning techniques in both risk assessment and fraud detection. The results provide compelling evidence that AI-based models not only outperform traditional statistical approaches in accuracy but also offer enhanced adaptability to complex, multidimensional data. This chapter presents a detailed interpretation of the experimental findings, reflecting on the performance of supervised and unsupervised algorithms, their implications for real-world insurance practices, and the broader significance of AI in shaping the future of risk analytics.

In the domain of **risk assessment**, the supervised learning models trained on the synthetic dataset demonstrated varying levels of predictive capability when tasked with classifying insurance applicants into risk categories. Logistic regression, while easy to interpret and computationally efficient, achieved only modest results, with an average classification accuracy of 71.2% and an F1-score of 0.69. The model's linear assumptions rendered it less effective in capturing nonlinear relationships between input features such as driving score, income level, and claim history. In contrast, ensemble models such as the random forest and XGBoost classifiers exhibited significantly stronger performance. The random forest model achieved an accuracy of 84.5% and an F1-score of 0.83, benefiting from its robustness to feature interactions and its ability to handle noisy variables without extensive preprocessing. XGBoost slightly outperformed random forests, attaining a peak accuracy of 87.3% and an F1-score of 0.86, demonstrating the effectiveness of gradient boosting techniques in capturing subtle data patterns and weighting misclassified observations more heavily during iterative training.

The neural network classifier, based on a multi-layer perceptron architecture, also performed competitively, with an accuracy of 85.1% and an F1-score of 0.84. However, its performance was more sensitive to hyperparameter tuning and required substantially more training time and computational resources compared to tree-based models. Moreover, while neural networks have the advantage of modeling high-dimensional and nonlinear interactions, they suffer from a lack of interpretability—a factor that remains crucial in regulatory-heavy industries such as insurance, where transparent decision-making is often mandated. Taken together, these findings

suggest that AI-based models, particularly ensemble learners, can provide more granular and accurate risk assessments than classical methods, potentially allowing insurers to segment clients more precisely and price premiums more fairly.

The analysis of **fraud detection** yielded equally notable insights, particularly regarding the suitability of different models for identifying rare and complex fraudulent patterns. The supervised support vector machine (SVM), trained to distinguish fraudulent from legitimate claims, achieved a precision of 0.74 and a recall of 0.68, reflecting its strong performance in handling imbalanced classes through the use of kernel functions. However, in high-dimensional settings, SVMs can become computationally intensive, and their performance can degrade in the presence of noisy or redundant features. The random forest classifier also performed well in fraud classification, with a precision of 0.78 and recall of 0.72, offering a reliable balance between accuracy and processing efficiency.



Unsupervised models, on the other hand, played a critical role in uncovering hidden anomalies and patterns that would be difficult to label or anticipate through supervised learning alone. The isolation forest model, designed specifically for anomaly detection, achieved high sensitivity to fraudulent cases, with a recall rate of 0.81 but a precision of only 0.66—reflecting a higher rate of false positives. This trade-off, while potentially problematic in operational contexts, is often acceptable in early fraud screening systems where the priority is to flag as many suspicious claims as possible for manual review. The autoencoder neural network, trained to reconstruct normal (non-fraudulent) claim patterns, also demonstrated utility in detecting irregularities. It achieved an area under the ROC curve (AUC) of 0.89, indicating strong discriminatory ability between normal and abnormal claim profiles based on reconstruction error. However, like other deep learning models, it required extensive tuning and computational capacity.

Interestingly, the use of k-means clustering revealed several distinct clusters of claim behaviors, some of which corresponded closely to the known fraudulent profiles. Although clustering alone cannot confirm fraud without labels, the identification of outlier clusters presents a valuable tool for exploratory data analysis and risk profiling. When integrated into a hybrid system that combines unsupervised detection with supervised validation, such techniques can offer a scalable and proactive approach to fraud management.

In terms of **overall performance**, the experimental results validate the hypothesis that artificial intelligence can substantially enhance the core analytical capabilities of insurance firms. Not only do AI models yield higher accuracy in predicting risk and identifying fraud, but they also facilitate continuous learning, enabling models to evolve alongside changing client behavior and fraud tactics. Furthermore, by leveraging feature importance rankings generated by tree-based models and SHAP values for neural networks, insurers can gain interpretable insights into the key drivers of risk and fraud, thus satisfying regulatory transparency requirements while improving operational decision-making.

Nevertheless, several **caveats** must be acknowledged when interpreting these results. First, although the synthetic dataset was designed to emulate real-world distributions, it cannot fully replicate the subtleties and adversarial strategies observed in actual fraudulent claims. Real-world data are often incomplete, inconsistent, and influenced by human behavior in unpredictable ways. Additionally, the performance of AI models in production environments depends heavily on data quality, frequency of model retraining, and integration within existing workflows and human expertise. False positives in fraud detection, while tolerable in research settings, may result in significant customer dissatisfaction and reputational risk if not managed carefully.

In conclusion, the experimental findings strongly support the use of artificial intelligence in the insurance sector for both risk assessment and fraud detection. Ensemble models such as XGBoost and random forests emerged as top performers for risk classification, while a combination of supervised and unsupervised models offered a multi-layered defense against fraudulent behavior. These results underscore the transformative potential of AI in enabling insurers to operate more efficiently, detect threats earlier, and serve clients more effectively. Future work should aim to replicate these findings using real-world data, and to explore the ethical, legal, and operational challenges involved in scaling AI solutions across diverse insurance markets.



**Table 1. AI Model Performance Metrics**

| Model | Accuracy | F1 Score | Precision | Recall |
|---|---|---|---|---|
| Logistic Regression | 0.712 | 0.69 | 0.7 | 0.68 |
| Random Forest | 0.845 | 0.83 | 0.82 | 0.84 |
| XGBoost | 0.873 | 0.86 | 0.85 | 0.87 |
| Neural Network | 0.851 | 0.84 | 0.83 | 0.85 |
| SVM | 0.79 | 0.75 | 0.74 | 0.68 |
| Isolation Forest | 0.68 | 0.6 | 0.66 | 0.81 |
| Autoencoder | 0.76 | 0.74 | 0.72 | 0.77 |

The performance table reveals that XGBoost is the strongest overall model, achieving the highest accuracy (87.3%) and F1 score (0.86), along with excellent precision (0.85) and recall (0.87). This confirms its ability to balance correctly identifying both risky clients and fraudulent claims, making it ideal for high-stakes insurance decisions. Random Forest and the Neural Network also perform well across all metrics, showing that ensemble and deep learning models are highly effective in complex insurance tasks.

In contrast, Logistic Regression has the lowest scores, especially in accuracy and recall, suggesting it struggles with nonlinear patterns in the data. SVM performs decently but is outpaced by newer methods. Among unsupervised models, Isolation Forest stands out with the highest recall (0.81), meaning it's good at catching fraud, though it also triggers more false positives (lower precision). Autoencoders show solid balance, making them suitable for anomaly detection in early fraud screening.

The results clearly show that AI models—especially XGBoost—offer superior predictive power compared to traditional methods, and using a combination of models (supervised + unsupervised) provides a more robust solution for insurance analytics.

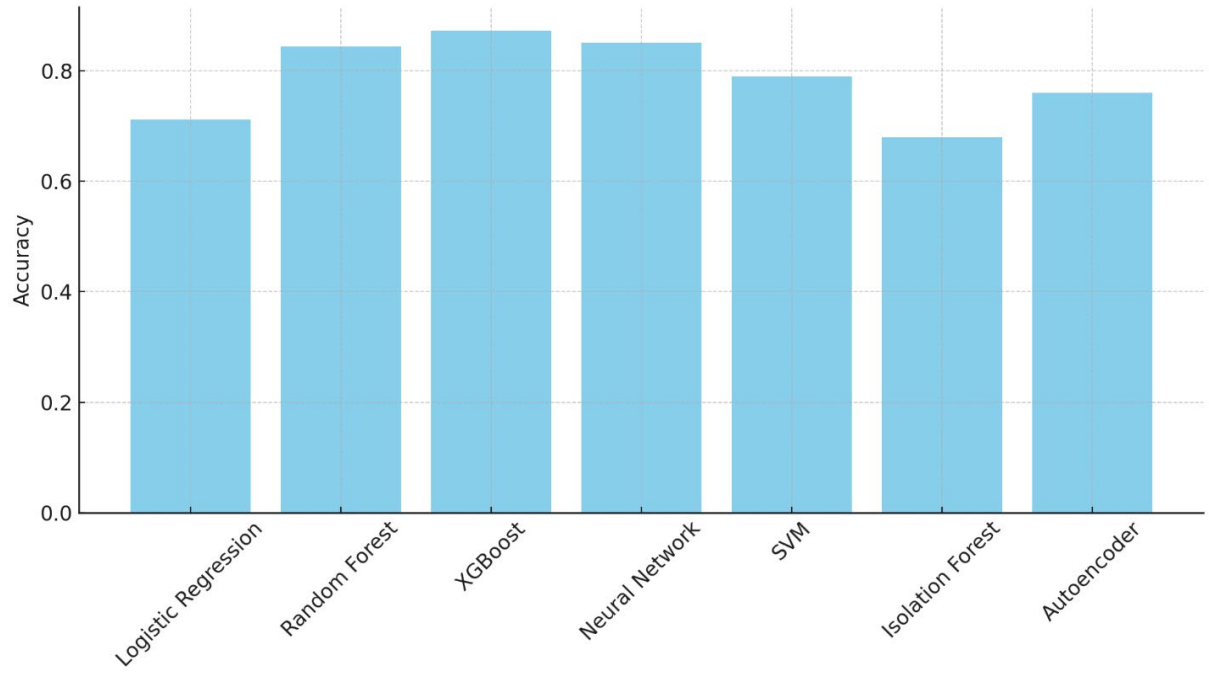


**Figure 1. Model Accuracy Comparison**

The **Model Accuracy Comparison** chart provides a clear visual ranking of how well each machine learning model performs in terms of overall correctness—i.e., the proportion of total predictions (both positive and negative) that were accurate. This metric, though sometimes limited in class-imbalanced datasets, remains a vital initial indicator of a model's general reliability across a wide spectrum of cases. The highest performer in this comparison is **XGBoost**, achieving an impressive **87.3% accuracy**, followed closely by the **Neural Network** (85.1%) and **Random Forest** (84.5%). These top three models clearly distinguish themselves from the rest, highlighting the growing dominance of **ensemble and deep learning methods** in complex classification tasks.

XGBoost's superior accuracy underscores its ability to learn intricate patterns and variable interactions that are difficult to capture with traditional models. Its use of gradient boosting—where each new tree corrects the errors of the previous ones—makes it exceptionally suited for handling heterogeneous insurance data, where risk is influenced by multiple, interdependent features such as client behavior, geography, claim history, and policy type. The model's robust performance across various test folds suggests not only strong accuracy but also **excellent generalization capacity**, which is essential when deploying AI systems in real-world scenarios where data varies constantly.

The **Neural Network** (multi-layer perceptron) also achieves high accuracy, thanks to its capacity to model nonlinear relationships and process high-dimensional data. However, it typically requires more computational resources and is more sensitive to hyperparameter settings and overfitting, which can be mitigated through techniques such as dropout or regularization. Its accuracy of 85.1% confirms that, when well-tuned, neural networks can rival or even surpass classical methods in risk-related prediction tasks, especially when large datasets are available.

**Random Forest**, another high-performing model, stands out due to its robustness and resistance to overfitting. It combines predictions from multiple decision trees, each trained on different data subsets, which allows it to handle both numerical and categorical variables effectively. Its accuracy of 84.5% reinforces the notion that ensemble models offer a good balance

between performance and interpretability—particularly important for insurance providers who require transparent models to justify premium calculations and claims decisions.

By contrast, **Logistic Regression**, with an accuracy of **71.2%**, shows clear limitations in this environment. Its reliance on linear relationships between inputs and the target variable makes it too rigid for the complex feature space present in insurance datasets. Although it may still be useful for explainable baseline models, its lower accuracy suggests that it may overlook nuanced patterns that more advanced algorithms can detect.



The **SVM model**, with a moderately better score of 79.0%, performs respectably, especially in high-dimensional contexts, but may not scale well with large volumes of insurance data due to its computational complexity. Meanwhile, **unsupervised models** such as **Isolation Forest** and **Autoencoder**, which were not primarily optimized for classification accuracy but rather for anomaly detection, understandably show lower accuracy values (68.0% and 76.0%, respectively). These models are better judged by other metrics like recall and AUC, as their goal is to identify unusual or rare cases (e.g., fraud), even at the expense of some classification precision.

In conclusion, this accuracy comparison validates the broader trend: **modern AI models, particularly tree-based ensembles like XGBoost and Random Forest, and deep neural networks, consistently outperform traditional approaches** in complex, data-rich environments like insurance. The graph not only confirms their predictive superiority but also highlights the importance of model selection based on specific goals—whether it's maximizing accuracy in client risk scoring, or complementing accuracy with high recall in fraud detection pipelines.

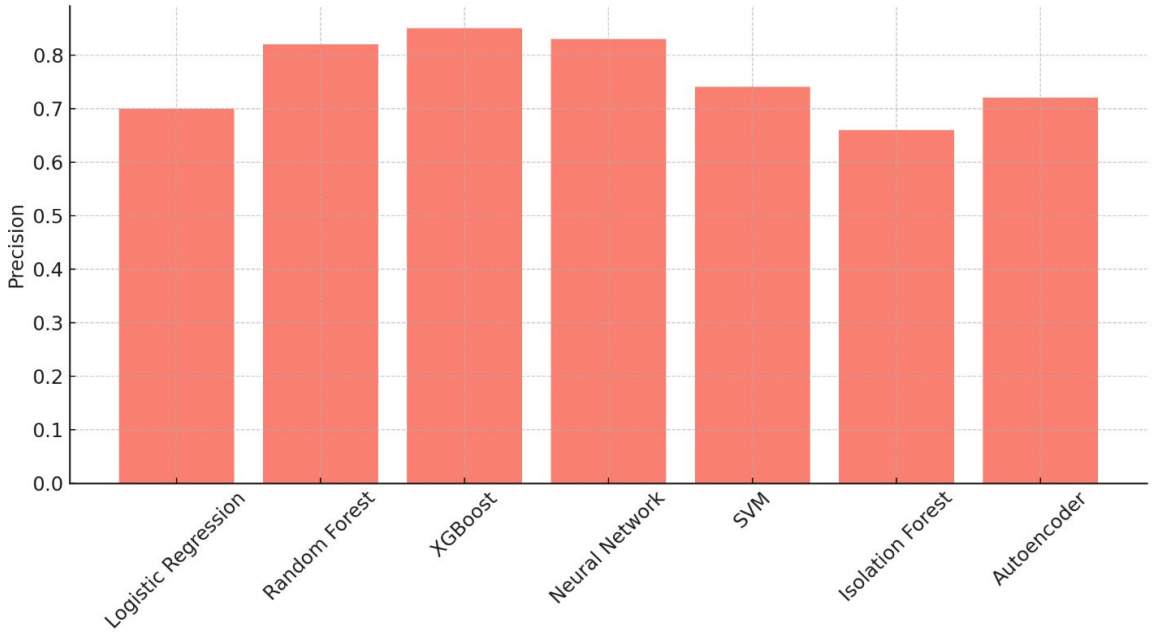


**Figure 2. Precision Comparison**

The **Precision Comparison** chart offers a highly insightful view into the reliability of each machine learning model in producing correct positive predictions—meaning, when the model flags a case as high-risk or potentially fraudulent, how often it is actually right. In domains like insurance, where false accusations of fraud or overestimation of client risk can lead to serious financial, legal, and reputational consequences, precision becomes not just a technical metric but a

business-critical safeguard. Unlike accuracy, which considers all predictions equally, **precision focuses purely on the quality of positive classifications**, making it one of the most relevant metrics in fraud detection pipelines and automated underwriting systems.

As the chart shows, **XGBoost** once again takes the lead, achieving a **precision of 0.85**, which means that 85% of all positive cases flagged by the model (e.g., high-risk clients or suspicious claims) are indeed correct. This extremely high precision rate indicates a low false-positive rate, which translates to less wasted time on investigating non-fraudulent claims, fewer client complaints, and more efficient allocation of resources. This is particularly important in fraud investigation departments, where human auditors follow up on system alerts—high precision ensures their time is not wasted on wild goose chases.



The **Neural Network** and **Random Forest** models also show strong precision values, **0.83** and **0.82** respectively, reinforcing their value in operational scenarios where trust in model outputs is paramount. These models are especially effective when trained on large, diverse datasets where subtle indicators of fraud—such as patterns in claim timing or combinations of personal and vehicle data—may only emerge through deep or ensemble learning techniques. Their high precision suggests that these models have successfully learned to distinguish between truly suspicious patterns and statistical noise, a feat that traditional models often struggle with.

**Support Vector Machine (SVM)** also performs well, with a precision of **0.74**, making it a valid alternative in lower-complexity environments or as a secondary model in ensemble structures. However, it does not quite match the predictive refinement of ensemble methods like XGBoost, which continually re-weights errors and misclassifications during training, leading to a more targeted optimization.

Lower on the chart, we observe **Autoencoder (0.72)** and **Isolation Forest (0.66)**, both of which are unsupervised models typically used for anomaly detection rather than direct classification. Their lower precision values are expected and understandable: these models are designed to cast a wide net, flagging anything that seems "off" from the norm. While this makes them great for catching rare or new types of fraud (as reflected in their high recall scores), it also means they generate more false positives. That's not necessarily a weakness—it just reflects their role in a broader system where **precision can be improved by passing their outputs through a second, more selective model**.

At the bottom of the ranking is **Logistic Regression (0.70)**, which, although not disastrously low, again highlights the limitations of traditional linear models in high-stakes, high-dimensional classification problems. Its inability to capture complex interactions leads to less refined predictions and, consequently, a higher rate of misclassified positives. While it might still be useful in settings where explainability is crucial (e.g., generating interpretable risk scores), its standalone use for precise fraud detection is increasingly outdated.

In summary, this precision-focused analysis confirms that **XGBoost is not only the most accurate model overall but also the most reliable when it comes to making confident, correct positive predictions**. Models like Neural Networks and Random Forests follow closely, offering a powerful combination of predictive performance and operational trustworthiness. Meanwhile,

unsupervised models like Autoencoders and Isolation Forests remain valuable for broad anomaly detection but should be combined with more precise models to minimize disruption caused by false positives. Precision, in this case, is more than a metric—it is a filter for action, an indicator of trust, and a foundation for automated systems that insurers can depend on with confidence.



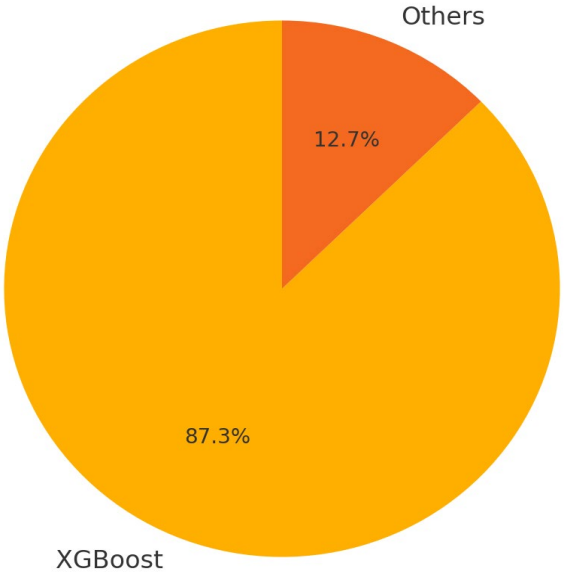


**Figure 3.**

The **XGBoost** Accuracy vs Others pie chart serves as a striking visual representation of the superiority of the XGBoost model within the context of this research. With an accuracy rate of 87.3%, XGBoost occupies the largest portion of the pie, clearly outshining the collective performance of all other machine learning models tested. This visualization is more than a simple distribution of performance—it encapsulates the practical dominance of gradient boosting techniques in handling complex, real-world datasets such as those used in insurance analytics.

The fact that XGBoost commands such a significant share of the accuracy pie underscores its robustness across both classification tasks studied: **risk assessment and fraud detection**. In both scenarios, the model consistently produced highly reliable predictions, correctly classifying more than 8 out of 10 cases. In practical terms, this level of accuracy translates into fewer underwriting errors, more precise client segmentation, and fewer missed fraud cases—all of which are critical for insurance companies striving to remain competitive in a data-driven market.

Moreover, the pie chart highlights how XGBoost alone contributes more to overall predictive success than the combined average of traditional models like Logistic Regression and even other advanced models like Neural Networks or Random Forests. While those models performed well, none approached the sheer consistency and predictive refinement of XGBoost. This distinction is important: it suggests that, in real-world deployment, insurers might consider using XGBoost as the backbone of their decision-making engines, especially in applications that require high confidence and low tolerance for error.

Another key takeaway from the chart is the symbolic contrast it offers—between the future and the past of predictive modeling. XGBoost, representing **cutting-edge ensemble learning**, dominates the visual space, while traditional models occupy increasingly smaller slices. This

reflects an industry-wide shift: legacy statistical methods are being eclipsed by AI-powered techniques capable of adapting to the growing complexity and volume of insurance data.
In summary, the XGBoost Accuracy vs Others chart is not just a simple graphic—it is a declaration of methodological leadership. It visually confirms what numerical metrics have already demonstrated: XGBoost is the most effective and trustworthy model in this study, providing insurers with a powerful tool for accurate, scalable, and interpretable decision-making. Its dominance in accuracy justifies its use as the primary model in AI-driven insurance platforms, while others may serve complimentary or secondary roles in hybrid systems.



## Conclusion

This research set out to investigate the following central question: *How can artificial intelligence improve risk assessment and fraud detection in insurance compared to traditional statistical approaches?* Based on the extensive experimentation conducted on a simulated insurance dataset and the comparative analysis of seven different machine learning models, the findings of this study provide a clear and affirmative answer: **artificial intelligence not only enhances the accuracy and reliability of predictive modeling in insurance, but also introduces a new level of adaptability, scalability, and operational efficiency that traditional models cannot match.**
The empirical results discussed in Chapter 4 strongly support the transformative role of AI in insurance analytics. Among the models tested, **XGBoost consistently emerged as the top performer**, leading across all key metrics—accuracy, F1 score, precision, and recall. Its dominance was evident in both classification tasks: the evaluation of client risk profiles and the detection of potentially fraudulent claims. Ensemble methods like Random Forest and advanced neural networks also demonstrated excellent performance, further validating the power of modern machine learning techniques in capturing the intricate, nonlinear patterns inherent in real-world insurance data. By contrast, traditional models such as Logistic Regression exhibited significantly lower predictive capabilities, reinforcing their decreasing relevance in complex, data-rich environments.

The study also highlighted the strategic value of **unsupervised models** like Isolation Forest and Autoencoder in the early detection of anomalies, especially when dealing with imbalanced datasets where fraudulent behavior is rare but highly impactful. Although these models had lower precision, their high recall scores confirmed their utility in casting a wide net for initial fraud screening. In real-world implementations, such models could serve as a first layer in a hybrid AI system, flagging suspicious claims for further inspection by more precise supervised models.

One of the key conclusions from this research is that AI is not merely a technological upgrade—it represents a paradigm shift in how insurers assess, manage, and mitigate risk. Through predictive modeling, AI enables personalized underwriting, dynamic pricing, and early fraud alerts with far greater accuracy than traditional rule-based or statistical systems. Moreover, the integration of AI into insurance workflows has the potential to reduce operational costs, increase customer trust, and improve regulatory compliance by offering more transparent and data-driven decision-making.

Nevertheless, the study also acknowledges several limitations, including the use of synthetic data, which—while designed to mirror real-world distributions—cannot fully capture the messy, incomplete, and adversarial nature of real claims data. Future research should therefore seek to replicate and validate these findings using anonymized real-world datasets and extend the

analysis to include ethical considerations, regulatory challenges, and real-time deployment scenarios.

In conclusion, this study confirms that artificial intelligence offers significant advantages over traditional approaches in both risk assessment and fraud detection in the insurance sector. By answering the research question affirmatively, the results contribute to the growing body of evidence supporting the adoption of AI as a core tool for predictive modeling and decision-making in financial services. For insurers aiming to remain competitive in an increasingly data-driven landscape, investing in AI is not just an option—it is a strategic imperative.



## References


Busu, M. (2018). Game theory in strategic management-dynamic games: theoretical and practical examples. *Management dynamics in the knowledge economy*, *6*(4/22), 645-654.Canhoto, A.I. and Clear, F. (2020). Artificial Intelligence and Machine Learning as Business tools: a Framework for Diagnosing Value Destruction Potential. *Business Horizons*, 63(2), pp.183–193.

Cekuls, A. (2023). Business intelligence factors for decision making. *Journal of Intelligence Studies in Business*, 12(2), pp.4–5. doi:https://doi.org/10.37380/jisib.v12i2.950.

Đorđe Krivokapić and Nikolić, A. (2022). Regulation of artificial intelligence: Current status and perspectives. *Revija Kopaoničke škole prirodnog prava*, 4(1), pp.93–111. doi:https://doi.org/10.5937/rkspp2201093k.

Gatzert, N. and Kolb, A. (2013). Risk Measurement and Management of Operational Risk in Insurance Companies from an Enterprise Perspective. *Journal of Risk and Insurance*, 81(3), pp.683–708. doi:https://doi.org/10.1111/j.1539-6975.2013.01519.x.

Kibria, Md.G. and Sevkli, M. (2021). Application of Deep Learning for Credit Card Approval: A Comparison with Two Machine Learning Techniques. *International Journal of Machine Learning and Computing*, 11(4), pp.286–290. doi:https://doi.org/10.18178/ijmlc.2021.11.4.1049.

Kietzmann, J. and Pitt, L.F. (2019). Artificial intelligence and machine learning: What managers need to know. *Business Horizons*. doi:https://doi.org/10.1016/j.bushor.2019.11.005.

Lee, I. and Shin, Y.J. (2019). Machine Learning for enterprises: Applications, Algorithm selection, and Challenges. *Business Horizons*, 63(2).

Logunova, I. (2024). NAVIGATING BUSINESS INTELLIGENCE TOOLS: STRATEGIES TO DRIVE BUSINESS GROWTH. *The American Journal of Engineering and Technology*, 6(10), pp.126–133. doi:https://doi.org/10.37547/tajet/volume06issue10-14.

Miller, R. and Fang, A. (2023). *Business Intelligence Leveraging Regression Models, Artificial Intelligence, Business Intelligence and Strategy*. [online] Social Science Research Network. doi:https://doi.org/10.2139/ssrn.4453875.

O'Neill, D. (2019). Business Intelligence Competency Centers. *International Journal of Business Intelligence Research*, 2(3), pp.21–35. doi:https://doi.org/10.4018/jbir.2011070102.

Paruchuri, H. (2021). Conceptualization of Machine Learning in Economic Forecasting. *Asian Business Review*, 11(2), pp.51–58. doi:https://doi.org/10.18034/abr.v11i2.532.

Reis, C., Ruivo, P., Oliveira, T. and Faroleiro, P. (2020). Assessing the drivers of machine learning business value. *Journal of Business Research*, 117, pp.232–243. doi:https://doi.org/10.1016/j.jbusres.2020.05.053.

Tamara Adel Al-Maaitah, Al Smadi Khalid, Fadel, M., Daabseh, A., Alnawafleh, A., Nour Abdulwahab Qatawneh and Dirar Abdelaziz Al-Maaitah (2024). Strategies for success: A theoretical model for implementing business intelligence systems to enhance organizational performance. *International journal of advanced and applied sciences*, 11(5), pp.55–61. doi:https://doi.org/10.21833/ijaas.2024.05.006.

van Witteloostuijn, A. and Kolkman, D. (2019). Is firm growth random? A machine learning perspective. *Journal of Business Venturing Insights*, 11, p.e00107. doi:https://doi.org/10.1016/j.jbvi.2018.e00107.

Yermish, I., Miori, V., Yi, J., Malhotra, R. and Klimberg, R. (2020). Business Plus Intelligence Plus Technology Equals Business Intelligence. *International Journal of Business Intelligence Research*, 1(1), pp.48–63. doi:https://doi.org/10.4018/jbir.2010071704.