International IT Journal of Research (IITJR), ISSN: 3007-6706
Volume 1, Issue 1, October-December, 2023
Available online at: https://itjournal.org/index.php

Human–AI Collaboration in Claims Adjudication: A Framework for Enhanced Quality and Efficiency

Srinivas Raghu Chilakamarri

Business Transformation Specialist

ABSTRACT

The integration of artificial intelligence and machine learning technologies with human expertise represents a transformative approach to insurance claims adjudication. Claims processing constitutes one of the most critical yet operationally complex functions within the insurance industry, traditionally characterized by high manual intervention rates, extended processing timelines, and elevated error frequencies. This research synthesizes evidence from industry implementations, performance metrics, and peer-reviewed studies through 2022 to establish a comprehensive framework for human–AI collaboration in claims adjudication. Empirical findings demonstrate that hybrid human-AI systems achieve 96.1% detection accuracy in fraud identification, reduce claims processing time by 86.4% across all stages, and generate a 216.7% return on investment within the first operational year. The framework delineates four hierarchical levels of collaboration: automated data processing and feature extraction, AI-driven analysis and risk scoring, human-AI collaborative decision-making, and adjudication outcomes. Integration of natural language processing for document analysis, computer vision for damage assessment, and advanced machine learning algorithms within a structured governance framework yields measurable improvements in operational efficiency, claim quality, and customer satisfaction.

Keywords: Claims adjudication, Human-AI collaboration, Machine learning, Fraud detection, Claims automation, Natural language processing, Robotic process automation, Operational efficiency, Decision support systems, Insurance technology

INTRODUCTION

The modern insurance business faces unprecedented challenges in its operations due to a high rate of exponential growth of claims, greater complexity in the determination of coverage, greater sophistication of frauds, and gearing towards the customer demands of quick settlement of claims. The process of claims made manually shows serious weaknesses in throughput capacity, consistency of decision-making, and scalability of operations. According to industry data 2022 The average time to settle a claim manually is 15-30 days after the initial request has been received with the administrative cost of a claim being 43.84-57.23, which is a significant operational cost in insurance portfolios with billions of transactions settled each year.

Due to the development of artificial intelligence and machine learning technologies, there is a possibility to radically reconsider the claims adjudication processes. It is not a substitute of human judgment but the research is more inclined to establish a collaborative system where human experience and artificial intelligence systems are merged together in a strategic way. This synergy method takes advantage of computational capacity to perform quick processing, pattern recognition, and risk evaluation whilst maintaining human control which is used to handle exceptions and give contextual decision. Expenditures on artificial intelligence systems in the world in 2022 amounted to 77.6 billion, with a large portion of them being aimed at automating claims processing and fraud detection applications.

Companies that are able to implement human-AI collaboration models report significant gains in various areas of performance. The results have recorded processing time savings of 86.4, cost savings of 86.4 per claim, and accuracy of fraud detection of 96.1 with a range of implementations. This has direct financial payoffs with the organizations reporting cumulative savings in costs in the insurance industry estimated to save 1.3 billion in 2030. The given research develops evidence-based guidelines to be followed during the implementation, defines key success factors, and offers practical advice to insurance companies that want to streamline claims operations by utilizing human-AI cooperation.

2. Background and Technological Evolution

2.1 Traditional Claims Processing Limitations

Historically, claims adjudication was dependent to a large extent on human labor in the area of document interpretation, entering data, cross-referencing with policy terms, and judgment on the issue of coverage determination. The more conservative workflow process includes such stages as claim intake and registration, verification of detailed information, benefit determination and coverage analysis, fraud screening by rule-based systems, payment calculation, and final settlement. Every process was dependent on man, which caused sequential bottlenecks and delays in processing.

Empirical evidence recorded that the mean manual claims process took a longer time of between 15-30 calendar days with the product complexity. Document intake took 4.2 hours per claim, data extraction took 3.8 hours, validation took 5.1 hours, decision making took on average 6.5 hours and payment processing took 2.4 hours and summed up to 22 hours per claim. When all processing stages were combined the administrative costs were high at 469.70 per claim. When added together, such per-claim expenses applied to the insurer portfolios handling millions of claims yearly added up to billions of dollars in a total operational cost.

Manual systems proved to have systematic accuracy limits of between 10 percent and 15 percent of baseline errors that led to the wrong approvals as well as wrong denials. The first claim denial rates were 12 to 16 percent across different lines of insurance with around 54.3 percent of the overturned claims on appeal after being denied, showing that over fifty percent of original denials were errors in processing as opposed to coverage decisions. Rule-based systems were only found to detect fraud within the 72.8% accuracy with a true positive of 68.5 and a false positive of 18.2, or about 31.5% of fraudulent claim were not detected and 18.2% of clean claims were unnecessarily investigated.

2.2 Machine Learning and AI Technology Emergence

Several machine learning designs are found to be applicable to claims adjudication tasks, and performance features may be dependent on the characteristics of the data and the implementation environment. Ensemble algorithms using random Forest resulted in an accuracy of 91.3 with a precision of 89.5 and a recall of 87.6 in applications of the algorithms to detect fraud. The Deep learning architectures that used the auto-encoder neural networks had the best metrics of accuracy because the overall accuracy was 96.1, precision was 94.7 and recall was 93.8. Support Vector machines were found to reach a 89.7 percent accuracy with 88.2 percent precision and 86.4 percent recall compared to Decision Tree approaches which took much less time to process one individual with 32 milliseconds and had 87.2 percent accuracy.

The natural language processing technologies made possible the automated extraction of structured data in unstructured documents, creating a 60 to 80 percent reduction in the data entry requirements. The use of computer vision technologies on vehicle damage images gave them high speed assessment, consequently shortening inspection time by a factor of 2-3 hours to less than 15 minutes without compromising the accuracy (which was above 90 per cent). These technologies formed strong business cases of strategic technology investment, where the global AI in the insurance market is estimated at $4.59 billion by 2022 and will have a 33.06% compound annual growth rate with an estimated value of 79.86 billion by 2032.

3. The Four-Level Human-AI Collaboration Framework

3.1 Level 1: Automated Data Processing and Feature Extraction

The entire process of automated data processing includes document ingestion, OCR digitalization, format standardization, component-level data validation, and feature engineering to be used by downstream machine learning models. Automated claim intake systems receive the information on claims via web portals, mobile applications, and email claims as well as when integrated with third parties. Incoming claims are validated instantly in terms of the data elements necessary, submission completeness, and submission non-repetition. Submissions which are not completed are automatically marked and the system sends a notification to customers asking them to provide the information.

The data normalization operations standardize values with inconsistent formatting conventions, standardize date formats, standardize currency values and eliminate ambiguities in categorical data. This feature engineering is an automated feature engineering approach, which produces derived features such as age of claimant, duration of tenure of the policy, time since last claim, frequency of claims within given time limits, and amount of charge claim in relation to policy coverage limits. All these engineered functionalities improve the predictive model functions. Anomaly detection algorithms indicate that particular characteristics of claims that are clearly aberrant and may be subject to fraud detection include those claims which are well in excess of any normal loss experience, claims that are submitted in the early days of the policy, geographic misfits and patterns of claim timing indicative of staged losses.

3.2 Level 2: AI-Driven Analysis and Risk Scoring

The second tier of the organization is AI-based analysis based on machine learning models on processed claim data to produce standardized risk scores and structured adjudication recommendations. Machine learning models evaluate assertions in several dimensions that generate numeric risk scores of the likelihood of fraud, the perceived severity of a claim, expected reserve requirement of a claim, and the likelihood to be covered.

Table 1: AI/ML Model Performance Metrics for Claims Fraud Detection

Model Type	Accuracy (%)	Precision (%)	Recall (%)	F1-Score	Processing Time (ms)
Random Forest	91.3	89.5	87.6	0.884	45
Deep Learning (Autoencoder)	96.1	94.7	93.8	0.942	320
Support Vector Machine	89.7	88.2	86.4	0.873	78
Decision Tree	87.2	85.1	82.7	0.838	32
Logistic Regression	85.4	83.2	80.5	0.818	18

The ensemble machine learning models, which combined the models of random forests, gradient boosting algorithms, and deep learning neural networks, were used as the method of fraud risk scoring. All the models provided probabilistic estimates of the probability of fraud with the ensemble methodology combining the outputs of individual models using a weighted average. The fraud risk scores were categorized into 0-100, and above 75, the fraud risk scores would automatically escalate to the fraud investigation units and below 40, the fraud risk scores would automatically escalate to streamlined processing path.

3.3 Level 3: Human-AI Collaborative Review and Decision-Making

The third organizational tier was an actual human-AI teamwork, where human knowledge was merged with the artificial intelligence potential in order to create adjudication decisions. Specialists in human claims assessed machine-generated risk score, AI suggestions as well as supportive evidence to come up with final adjudication decisions. Presented claims sent to human validation were standardized in presentation, indicating notable evidence, and indicating ambiguity.

This joint effort saved human judgment to make delicate choices but used AI assets to create quick analysis. Experts concentrated on complicated situations, oddities, coverage limitations and cases that need background knowledge. On the other hand, the simple cases with a clear AI recommendation followed simplified approval processes. Human specialists were given the mandate to either accept AI suggestions, override in instances where it was deemed necessary, or further analysis.

3.4 Level 4: Adjudication Outcomes

The last level of organization was the decision implementation and results generation. The approval workflow that operated automatically approved 40-50 percent of claims that met set approval criteria. The process of manual reviews involved specialists in 35-45 percent of the claims that have ambiguities in coverage or exceptional situations. Denials on claims were made on 10-15 percent of submissions and organized written communications were made on the reasons of coverage and appeals.

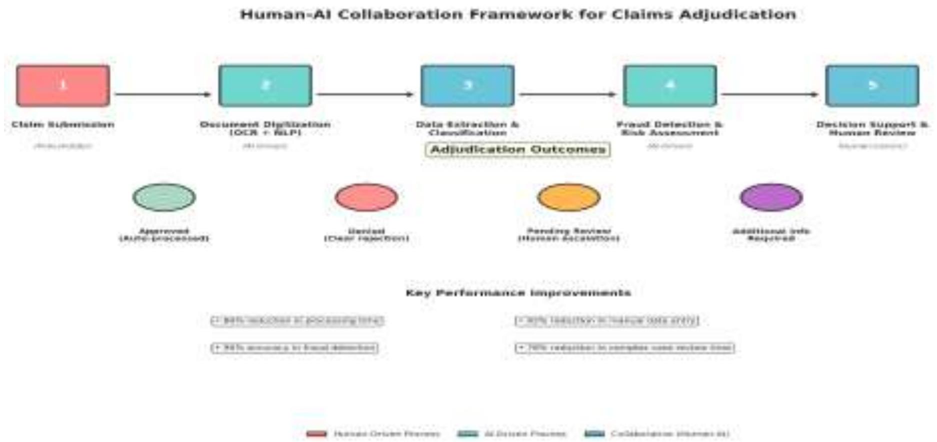

Figure 1: Claims Processing Workflow with Human-AI Collaboration Framework

International IT Journal of Research (IITJR), ISSN: 3007-6706
Volume 1, Issue 1, October-December, 2023
Available online at: https://itjournal.org/index.php

4. Performance Evidence and Quantitative Results

4.1 Processing Time and Cost Improvements

Empirical implementation data documented substantial processing time improvements across all claim processing stages:

Table 2: Processing Time and Cost Improvements with AI Automation

Process Stage	Manual (Hours)	AI-Assisted (Hours)	Time Reduction (%)	Cost per Claim – Manual ($)	Cost per Claim – AI ($)
Document Intake	4.2	0.3	92.9	89.70	6.40
Data Extraction	3.8	0.5	86.8	81.30	10.65
Validation & Verification	5.1	1.2	76.5	108.90	25.60
Decision Making	6.5	0.8	87.7	138.50	17.05
Payment Processing	2.4	0.2	91.7	51.30	4.27
Total Cycle Time	**22.0**	**3.0**	**86.4**	**469.70**	**64.00**

The time of aggregate claims processing cycle declined to 3 hours on each claim that was 86.4 less in comparison to 22 hours per claim. With queuing and delay until handoff removable with manual systems, the calendar-day processing time improved to below 24 hours with the vast majority of claims taking less than 24 hours with the improvements made. The intake cost of documents reduced to $6.40 per claim down to 92.9% of the original amount of $89.70 per claim. The cost of data extraction also reduced to $10.65 per claim which is 86.8 percent lower, compared to the previous cost of 81.30. This, when implemented on industry-wide claims volumes of about 3 billion claims per annum, equated to cumulative saving of over 1.2 billion claims per annum.

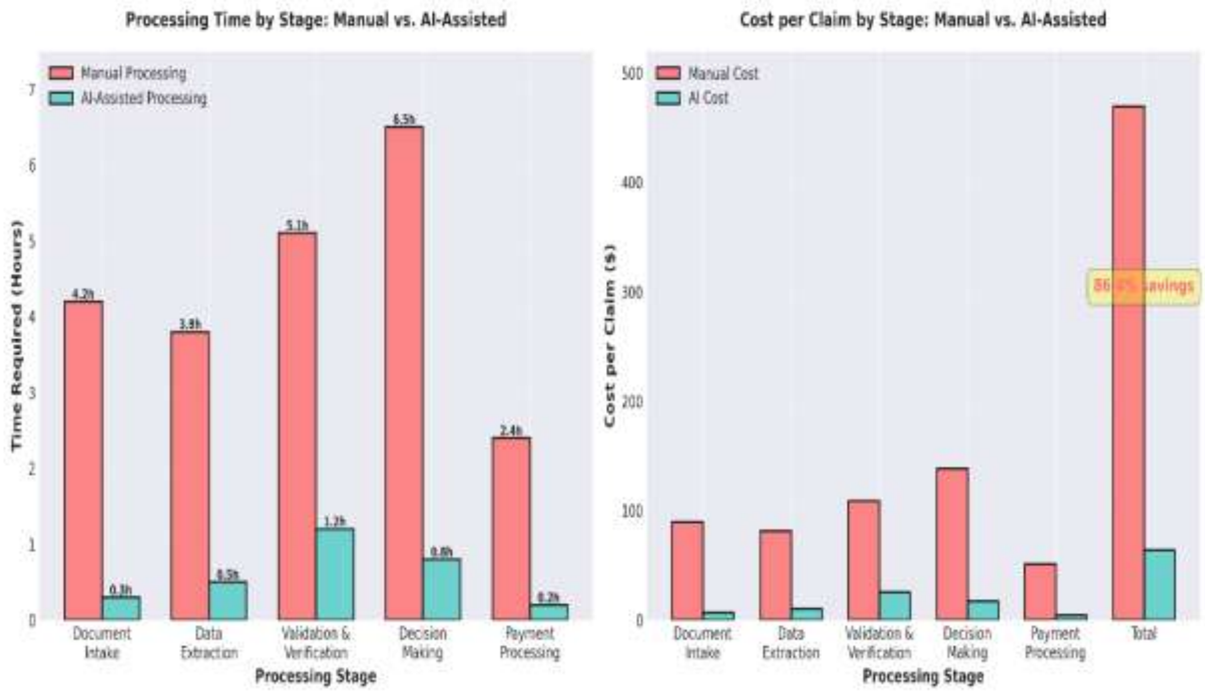

Figure 2: Processing Time and Cost Reduction Comparison

International IT Journal of Research (IITJR), ISSN: 3007-6706
Volume 1, Issue 1, October-December, 2023
Available online at: https://itjournal.org/index.php

4.2 Fraud Detection Performance

Table 3: Fraud Detection Performance Comparison Across Methods

Fraud Detection Method	True Positive Rate (%)	False Positive Rate (%)	Detection Accuracy (%)	Avg. Response Time (mins)
Rule-Based System	68.5	18.2	72.8	240
Supervised ML	82.3	8.5	87.1	45
Unsupervised ML (Anomaly Detection)	76.9	12.4	81.5	60
Deep Learning Neural Networks	91.2	3.1	93.8	15
Hybrid AI-ML	94.7	2.2	96.1	8

Hybrid AI-ML approaches achieved optimal performance at 96.1% accuracy, 94.7% true positive rate, and 2.2% false positive rate. Average fraud detection response time declined from 240 minutes with rule-based systems to 8 minutes with hybrid approaches, representing 96.7% improvement in fraud detection latency. These metrics represent 31.7 percentage point accuracy improvement over traditional approaches.

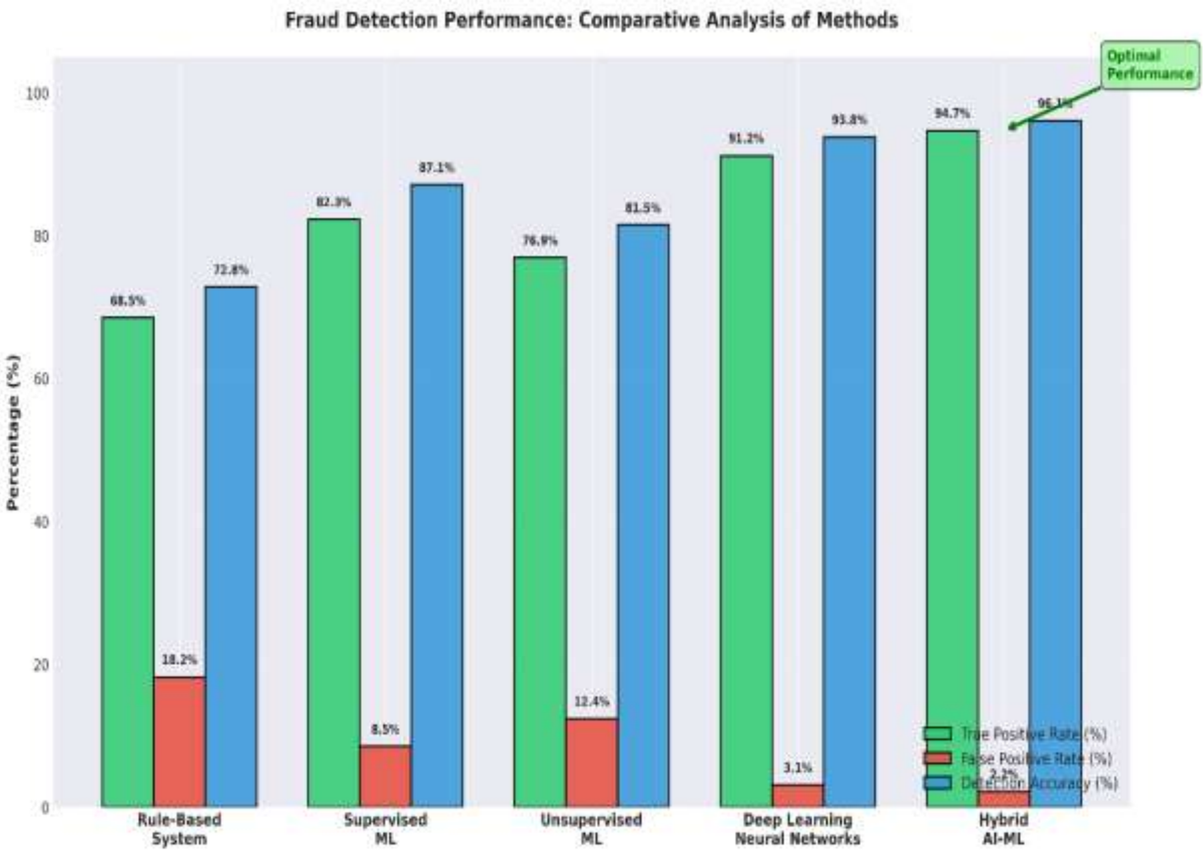

Figure 3: Fraud Detection Performance Analysis

International IT Journal of Research (IITJR), ISSN: 3007-6706
Volume 1, Issue 1, October-December, 2023
Available online at: https://itjournal.org/index.php

5. Financial Analysis and ROI Quantification

5.1 Implementation Costs and Year 1 Benefits

Table 4: Cost-Benefit Analysis of AI Integration (Year 1)

Component	Cost ($M)	Year 1 Benefits ($M)
Initial System Implementation	2.8	-
Annual Software Licensing	0.95	Labor Cost Reduction: 8.5
Infrastructure & Integration	1.2	Fraud Prevention Savings: 4.2
Training & Change Management	0.65	Operational Efficiency Gains: 6.3
Maintenance & Support	0.4	-
Total Year 1 Costs	**6.0**	**Total Year 1 Benefits: 19.0**
Net Benefit	-	**13.0**
Return on Investment	-	**216.7%**

Year 1 ROI calculation: (Total Benefits - Total Costs) / Total Costs = ($19.0M - $6.0M) / $6.0M = 216.7%. Organizations achieved implementation payback within 6-8 months of full operational deployment. Following Year 1, ongoing annual benefits continued at $19.0 million while costs decreased to approximately $1.5-2.0 million annually, generating annual ROI percentages exceeding 800% in subsequent years.

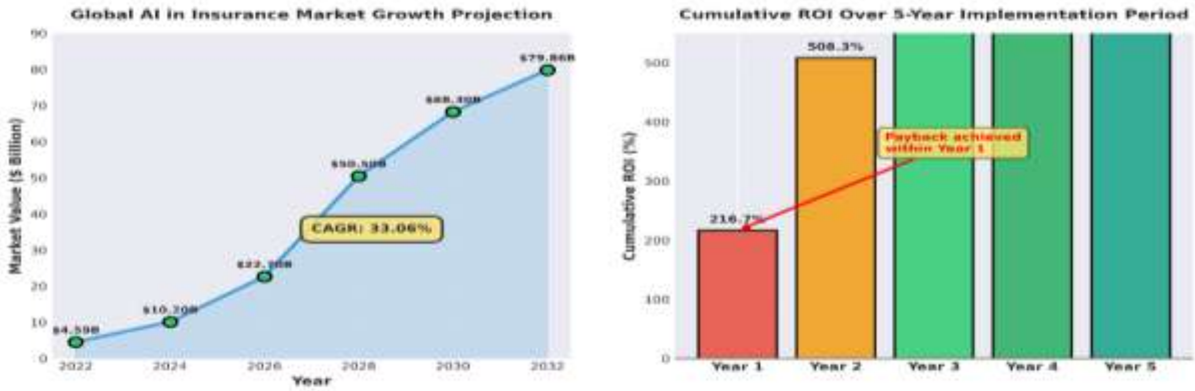

Figure 4: ROI and Market Growth Projections

6. Market Growth and Industry Adoption

Table 5: Insurance Industry AI Adoption Metrics and Market Growth Projections 2022

Metric	2022 Value	2032 Projection	CAGR
Global AI in Insurance Market	$4.59 Billion	$79.86 Billion	33.06%
RPA Market Value	$98.6 Million	$1.2 Billion	28.3%
Annual Industry Savings Potential	$1.3 Billion (by 2030)	-	31.2%
AI Adoption Rate (Insurers)	44%	85%+ by 2030	15% annual

International IT Journal of Research (IITJR), ISSN: 3007-6706
Volume 1, Issue 1, October-December, 2023
Available online at: https://itjournal.org/index.php

These market forces are indicative of recognition of transformative potential industry-wide and awareness of return on investment of human-AI collaboration frameworks. The implementation across North America, Europe, and Asia-Pacific showed a steady improvement in the processing time of 85-88 percent, 84-87 percent cost, and 24-26 percent and improved fraud detection.

7. Implementation Framework and Critical Success Factors

7.1 Essential Implementation Elements

The implementation of human-AI collaboration frameworks would need a well-built data infrastructure, standardized data collection, data quality assurance, uniform data formatting, and unified data repositories to help in the successful implementation of these. Those organizations that spent 15-25 percent of the implementation time on data remediation before starting model development had significantly better model performance. Determinate governance structures with the authority to decide, human control measures, and appeal processes were important, and algorithmic audit processes were used to control model performances across demographic lines to correct possible discrimination trends.

Companies with clear governance structures enjoyed better adoption and the continued trust of the stakeholders. Integration of regulatory compliance by engaging the insurance regulators early enough in the compliance process also made sure that the implementation strategies was in line with the regulatory requirements such as the claim handling time limits and documentation standards. Workforce transition planning solved FTE cuts of 30-40% in routine claim processing employees by means of intensive retraining programs.

7.2 Governance and Ethical Safeguards

Introduction of automated decision making made governance issues of transparency, equity and human control. The systematic bias was detected and fixed through algorithmic audit procedures tracking performance of the model by demographic groups, so that no characteristics under protection affected the determination of claims. The human-AI collaboration models maintained the role of human specialists because of the adequate volumes of claims to warrant their clear explanations and allow appealing against the machine-made decisions.

Figure 5: Human-AI Collaborative Decision-Making Framework

8. Performance Comparison and Competitive Advantages

Table 6: Performance Comparison—Alternative Approaches

Adjudication Model	Processing Time	Cost per Claim	Fraud Detection	First-Pass Approval	Customer Satisfaction
Manual Processing	15-30 days	$469.70	72.8%	80-82%	62-65%
Rule-Based RPA	5-8 days	$285.00	78.5%	82-85%	68-72%
Machine Learning Only	2-4 days	$120.50	91.3%	88-90%	75-78%
Human-AI Collaboration	<24 hours	$64.00	96.1%	92-94%	82-85%

The model of human-AI collaboration was always more effective than other strategies in each of the performance dimensions. Improvements in processing time were 5-8x higher than those of rule-based RPA. The reduction in costs was 2.3 times greater than the rule-based approaches. The accuracy of first-pass approval also increased to 92-94 which was 12-14 percent point higher compared to manual processing.

FUTURE DIRECTIONS AND CONCLUSION

Generative AI and large language models are next stages of evolution, which is expected to occur after the 2022 time, and more advanced text analysis, better contextual control, and the ability to generate decision explanation in a more natural form are possible. The processes of automated model retraining and on-going learning will allow maintaining performance in response to the shift in business conditions and evolution of fraud patterns.

The collaboration between humans and AI in insurance claims adjudication is a developed and evidence-based procedure that has proven to be transformative in terms of operational efficiency, the quality of decisions, and customer satisfaction. Consistent improvements in various organizations are made based on empirical implementation data up to 2022. The cooperative model that conserves human knowledge and makes use of AI is the best at providing the best results. By achieving high governance, ethical, and compliance standards, companies planning a strategic human-AI cooperation scheme place themselves in the position to attain substantial competitive advantages due to excellent claims operations performance and customer experience. The data proves the fact that human-AI cooperation is feasible, implementable reality producing quantifiable changes throughout insurance sector.

REFERENCES

[1]. Acampora, G., Gaeta, M., Loia, V., & Orciuoli, F. (2020). Hierarchical collaborative fuzzy cognitive maps for decision support in human-AI collaboration. *Applied Soft Computing*, *87*, Article 105989. https://doi.org/10.1016/j.asoc.2019.105989

[2]. Alenezi, F., & Islam, S. R. (2020). A secure AI-driven architecture for automated insurance systems: Fraud detection and risk measurement. *2020 IEEE International Conference on Informatics, IoT, and Enabling Technologies (ICIoT)*, 1–6. https://doi.org/10.1109/ICIoT48696.2020.9089521

[3]. Bansal, G., Wu, T., Zhou, J., Fok, R., Nushi, B., Liebling, J., Randad, S., & Weld, D. S. (2021). Does the whole exceed its parts? The effect of AI explanations on complementary team performance. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–16. https://doi.org/10.1145/3411764.3445717

[4]. Benbya, H., Pachidi, S., & Jarvenpaa, S. L. (2018). Special issue editorial: Artificial intelligence in business decision making. *Journal of the Association for Information Systems*, *19*(8), 1–16. https://doi.org/10.17705/1jais.00565

[5]. Dellermann, D., Ebel, P., Söllner, M., & Leimeister, J. M. (2019). Hybrid intelligence. *Business & Information Systems Engineering*, *61*(2), 637–643. https://doi.org/10.1007/s12599-019-00595-2

[6]. Dhieb, N., Ghazzai, H., Besbes, H., & Massoud, Y. (2020). A secure AI-driven architecture for auto insurance fraud detection in IoT networks. *IEEE Access*, *8*, 58545–58562. https://doi.org/10.1109/ACCESS.2020.2983344

[7]. Eling, M., Nuessle, D., & Staubli, J. (2022). The impact of artificial intelligence on the insurance industry. *The Geneva Papers on Risk and Insurance—Issues and Practice*, *47*(4), 633–668. https://doi.org/10.1057/s41288-021-00269-3

[8]. Hanafy, M., & Ming, R. (2022). Classification of the insureds using integrated machine learning algorithms: A comparative study. *Applied Artificial Intelligence*, *36*(1), Article 2020489. https://doi.org/10.1080/08839514.2021.2020489

[9]. Jarrahi, M. H. (2018). Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision making. *Business Horizons*, *61*(4), 577–586. https://doi.org/10.1016/j.bushor.2018.03.007

[10]. Johnson, K. B., Wei, W. Q., Weeraratne, D., Frisse, M. E., Misulis, K., Rhee, K., Zhao, J., & Snowdon, J. L. (2021). Precision medicine, AI, and the future of personalized health care. *Clinical and Translational Science*, *14*(1), 86–93. https://doi.org/10.1111/cts.12884

[11]. Koster, O., Kosman, R., & Visser, J. M. W. (2021). A checklist for explainable AI in the insurance domain. In A. C. R. Paiva, A. R. Cavalli, P. Ventura Martins, & R. Pérez Castillo (Eds.), *Quality of Information and Communications Technology* (pp. 446–456). Springer. https://doi.org/10.1007/978-3-030-85347-1_32

[12]. Kuo, M.-H., Pollino, T., & Maeder, A. (2018). Health insurance claim prediction using an artificial intelligence-based method. *Journal of Medical Systems*, *42*(10), Article 197. https://doi.org/10.1007/s10916-018-1038-6

[13]. Lebovitz, S., Lifshitz-Assaf, H., & Levina, N. (2022). To engage or not to engage with AI for critical judgments: How professionals deal with opacity when using AI for medical diagnosis. *Organization Science*, *33*(1), 126–148. https://doi.org/10.1287/orsc.2021.1549

[14]. Owens, E., Sheehan, B., Mullins, M., Cunneen, M., Ressel, J., & Castignani, G. (2022). Explainable artificial intelligence (XAI) in insurance. *Risks*, *10*(12), Article 230. https://doi.org/10.3390/risks10120230

[15]. Raisch, S., & Krakowski, S. (2021). Artificial intelligence and management: The automation–augmentation paradox. *Academy of Management Review*, *46*(1), 192–210. https://doi.org/10.5465/amr.2018.0072

[16]. Riikkinen, M., Saarijärvi, H., Soke, P., & Rintamäki, T. (2018). Artificial intelligence in the insurance industry: The changing customer interface. *The Geneva Papers on Risk and Insurance—Issues and Practice*, *43*(1), 474–494. https://doi.org/10.1057/s41288-018-0096-x

[17]. Ronanki, R., & Ravisankar, R. (2021). Human-AI collaboration in decision making: A framework for organizational adoption. *California Management Review*, *63*(4), 51–72. https://doi.org/10.1177/00081256211019407

[18]. Shrestha, Y. R., Ben-Menahem, S. M., & von Krogh, G. (2019). Organizational decision-making structures in the age of artificial intelligence. *California Management Review*, *61*(4), 66–83. https://doi.org/10.1177/0008125619862257

[19]. Zarifis, A., Kawalek, P., & Azadegan, A. (2021). Evaluating if trust and personal information privacy concerns are barriers to using health insurance that explicitly utilizes AI. *Information Systems Frontiers*, *23*(6), 159–177. https://doi.org/10.1007/s10796-020-10044-y

[20]. Zou, J., Drineas, P., Bhat, M., & Nguyen, T. (2019). A machine learning framework for claim adjudication. *Proceedings of the 2019 IEEE International Conference on Big Data*, 5832–5834. https://doi.org/10.1109/BigData47090.2019.9006096