Haoliang (Jack) Hu, Jenna Kim, Ethan Sonnenreich
MIS 373 IT Audit & Security
Dr. Hüseyin Tanriverdi
6 November 2023

Executive Summary: Analysis of Organizational Mitigation Mechanisms on AI Algorithm Risks

Introduction: In response to the growing demand for advanced risk management strategies in AI implementation, our consulting team embarked on a comprehensive analysis of various organizational mitigation mechanisms. This report encapsulates our findings on how these mechanisms influence the likelihood and impact of cybersecurity breaches, privacy infractions, and IT failures in AI algorithms.

Key Findings: *Statistical Analysis*: Cause-and-effect analyses, including correlation, logistic regression, and random forest models (the random forest model resulted in the best out-of-sample accuracy of 89%), pinpointed the most influential factors contributing to problem-free algorithms. These insights can steer organizations toward more effective risk management practices and are shared below. *Algorithm Usage*: Cognitive algorithms are predominant in the industry (94.68%). However, it was the lesser-used sensory and emotion-based algorithms (20.21%) that exhibited the fewest problems, suggesting a potential underutilization of safer AI technologies. *Learning and Optimization*: Firms that utilized supervised machine learning (only 29.79% of firms), engaged in multilateral optimization (46.81% of firms), and fostered inter-organizational collaboration (only 5.32% of firms) exhibited fewer issues. These practices appear to be pillars of a robust risk management framework. *Risk Mitigation Efforts*: A significant number of firms preferred symbolic (which makes up 22.34% of firms) over substantive risk mitigations (33.33% of firms). While symbolic measures, such as risk disclosures, were common, fewer firms invested in substantive measures, including oversight and design alterations, particularly for IT and privacy risks. However, having significant risk mitigation efforts proved to be an influential factor in causing algorithms to be problematic. *Risk Disclosure*: Cyber and privacy risk disclosures, at 70.21% and 84.04%, respectively, were notably higher than IT risk disclosures (at 55.85%). Less than half (48.94%) of firms comprehensively disclosed all three risk categories, even though the correlation and machine learning models showed having risk disclosures reduces the likelihood of being problematic. *Adverse Outcomes*: The dataset revealed instances of financial loss (17.02%), harm to individuals (12.76%), legal actions (6.38%), and algorithm discontinuation (4.26%) from problematic algorithms. These adverse outcomes underline the tangible consequences of insufficient risk mitigation.

Recommendations: *Expand the Use of Lower-Risk Algorithms*: Organizations can consider the broader implementation of sensory-based, emotion-based, and other low-risk AI technologies. *Invest in Supervised Learning and Optimization*: Companies should invest in supervised learning techniques and optimization strategies, which have demonstrated efficacy in minimizing AI risks. *Collaboration is Key*: Promoting inter-organizational collaboration can lead to shared learning and better risk management practices. *Prioritize Substantive Mitigations*: Firms must shift their focus from symbolic disclosures to substantive measures, integrating risk management into the design and oversight of AI systems. *Comprehensive Risk Disclosures*: We advocate for a holistic approach to risk disclosures, encompassing all facets of cybersecurity and privacy to foster transparency and preparedness.

Conclusion: The landscape of AI risks is complex and evolving. By adopting a strategic approach to risk mitigation, organizations can not only reduce the likelihood and impact of negative outcomes but also enhance the trustworthiness and reliability of AI systems. Our analysis provides a data-driven foundation for making informed decisions and positions our clients at the forefront of AI risk management.

Appendix

Exhibit A: Percentage of Firms with Each Algorithm Type

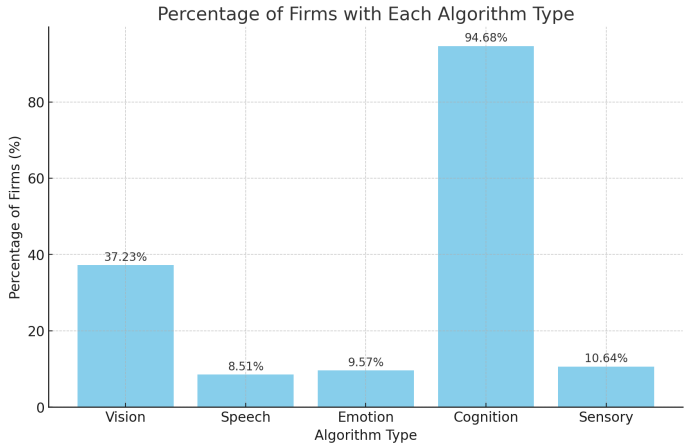


Exhibit B: Percentage of Firms Engaging in Symbolic vs. Substantive Risk Mitigations

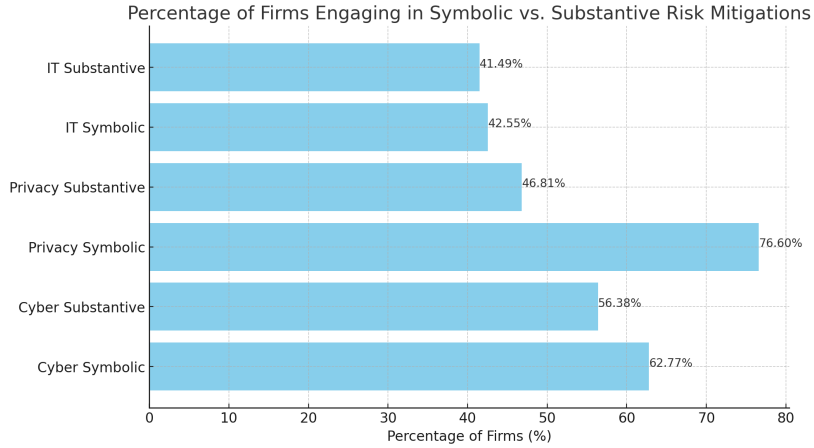


*Exhibit C: Descriptive Analyses of Group 3's Findings (Data From Regular Group Integrated Data)

Exhibit D: Count of Influential Factors to Target Variable

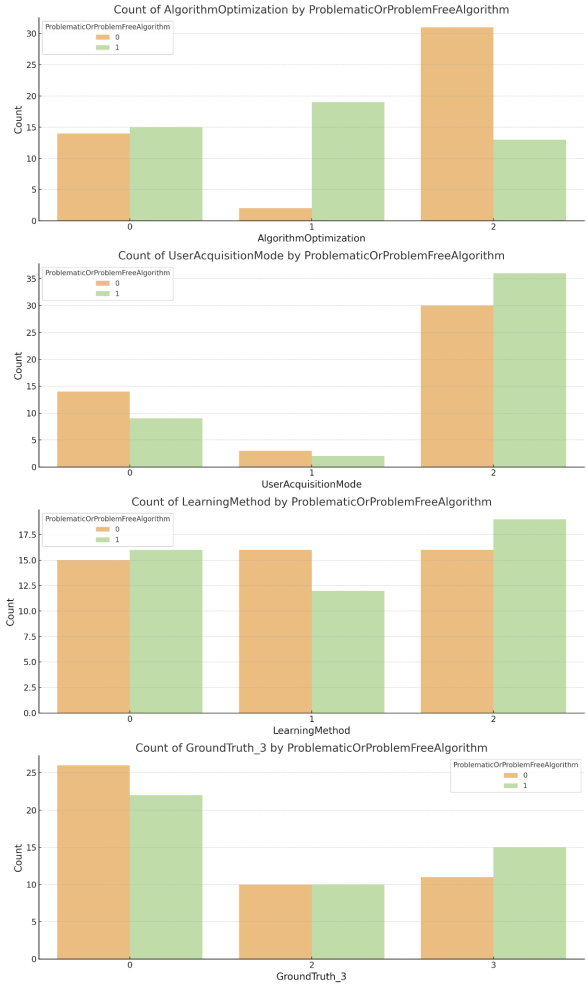


Exhibit E: Correlation to Outcomes Heatmap

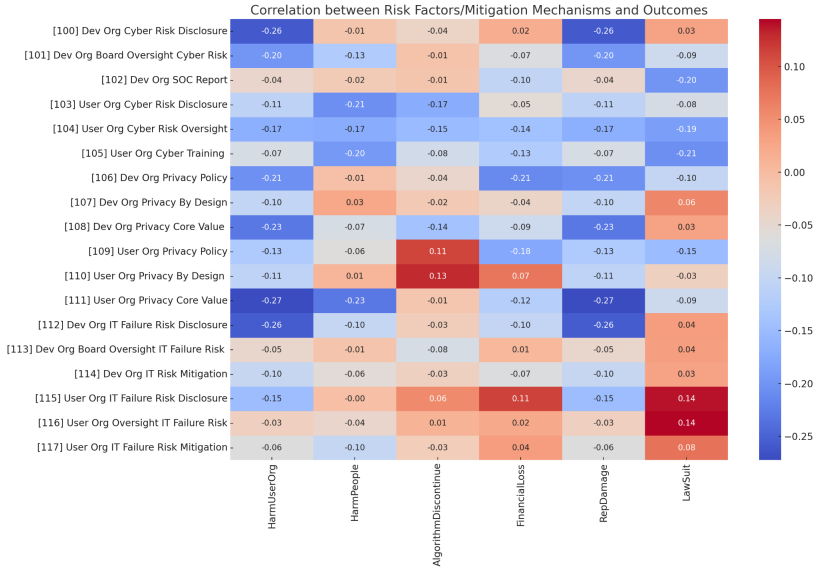


Exhibit F: Mean Values of Variables with Lowest p-values from Logistic Regression

Exhibit G: Percentage Distributions of Influential Factors

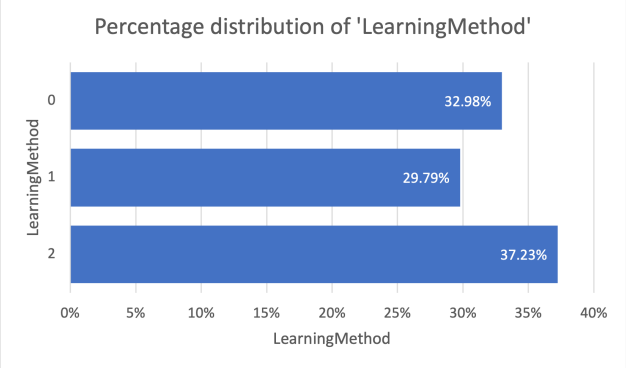
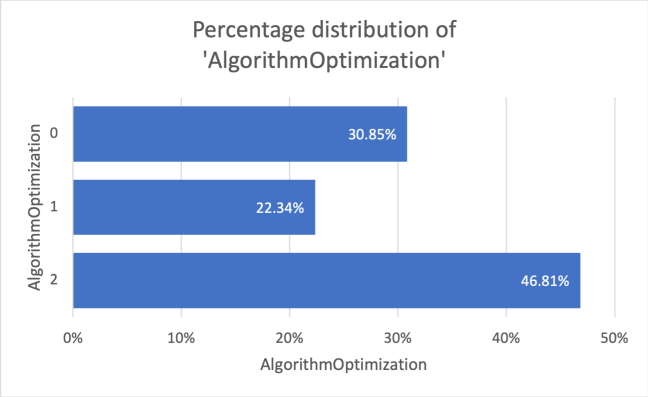


Exhibit G (continued)



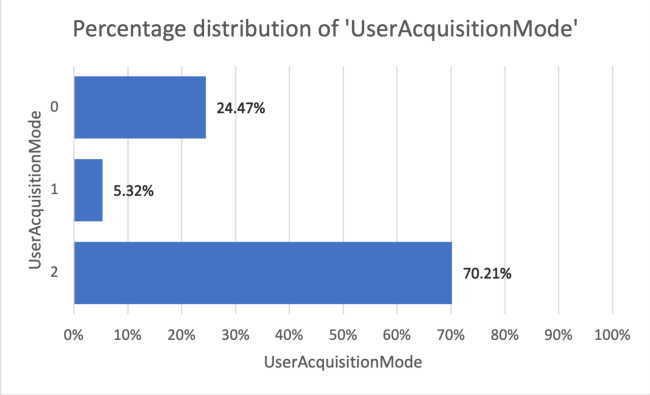


Exhibit H: Correlation Between the Predictors and the Target

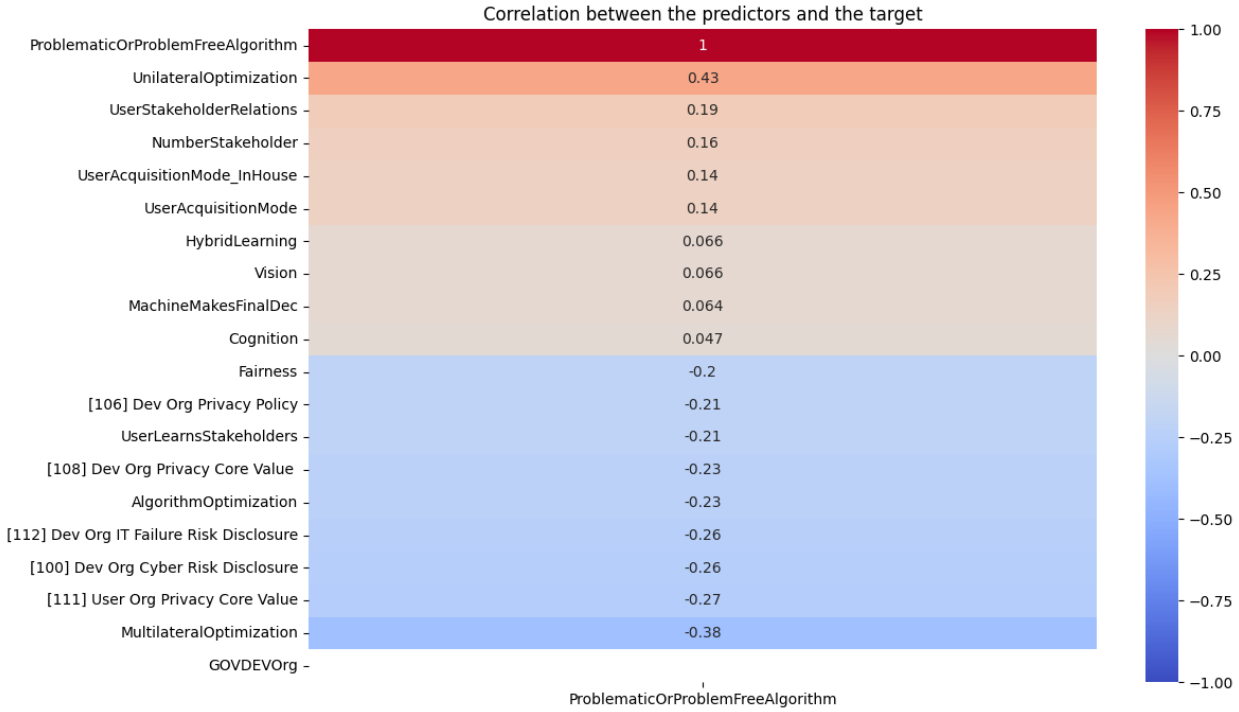


Exhibit I: Predictor Coefficients From Logistic Regression

	Coefficient
MultilateralOptimization	-1.159472
HumanMachineCollaboration	-0.962988
Accuracy	-0.562545
UserAcquisitionMode_Collaboration	-0.529096
Emotion	-0.526137
Fairness	-0.452300
[111] User Org Privacy Core Value	-0.432240
[112] Dev Org IT Failure Risk Disclosure	-0.383535
AlgorithmRepurposed	-0.318356
[100] Dev Org Cyber Risk Disclosure	-0.275138
MachineMakesFinalDec	-0.201714
Sensory	-0.167695
GOVUserOrg	-0.163223
MANUFUserOrg	-0.149773
Speech	-0.142254
SERVICEDEVOrg	-0.098080
OnPlatformOrOffPlatform	-0.090538
LearningMethod	-0.076610
[116] User Org Oversight IT Failure Risk	-0.054758
GOVDEVOrg	0.000000
[106] Dev Org Privacy Policy	0.046963
MANUFDEVOrg	0.098004
SERVICEUserOrg	0.312920
HybridLearning	0.333878
FidelityMatching	0.356372
NumberStakeholder	0.549299
UnilateralOptimization	0.832454
UserStakeholderRelations	0.984086

Exhibit J: Most Positive and Most Negative Coefficients

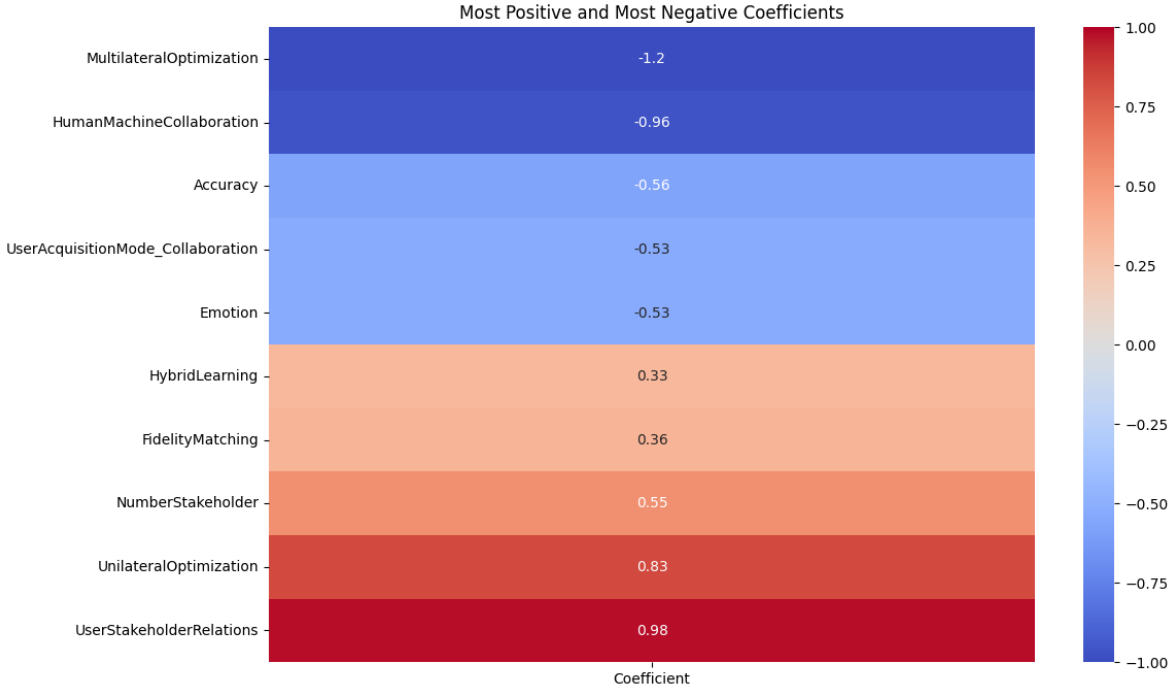
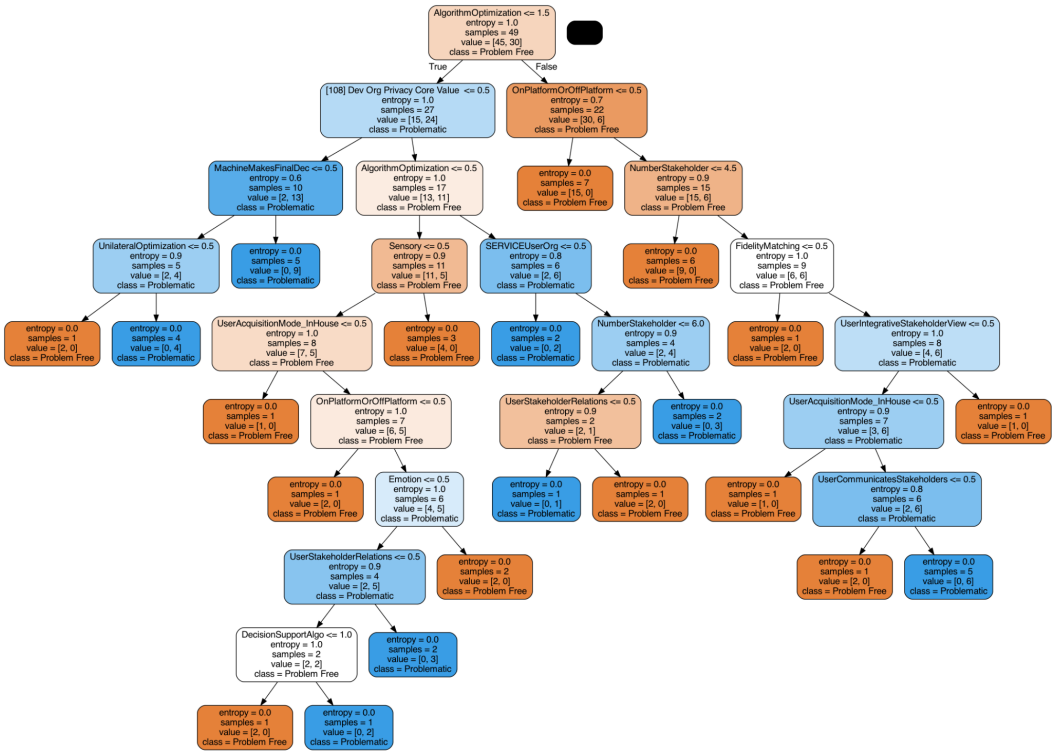


Exhibit K: Random Forest Classification



*Descriptive analyses were derived from Regular_Groups_Integrated_Data_10-30-2023.xlsx