Paper Review

AI and Corporate Risk Management: Identifying and Mitigating Technological and Ethical Risks

Problem Statement

Artificial Intelligence (AI) is rapidly being integrated across companies for different corporate tasks such as recruitment, credit scoring, customer service, supply-chain planning, quality control etc. Yet it creates new and different types of risks.

These risks generally fall into two broad groups:

- **Technological risks** - software bugs, model failures, data errors, sensor faults, security breaches.
- **Ethical risks** - unfair or biased outcomes, privacy violations, lack of transparency, and actions that harm reputation or stakeholder trust.

Companies often deploy AI without considering the consequences. This paper focuses on technological and ethical AI risks relevant to corporate contexts and also to assess how a unified risk management framework can detect, measure, and mitigate the risks.

Motivation

Companies use AI to screen thousands of job applications in minutes, to forecast supply chain demands, or to detect financial frauds. However, the rapid growth of AI also brings serious risks. The world has already seen cases where AI resulted in significant failures, such as:

- Recruitment tool that discriminates female applicants.
- Credit-scoring systems that is unfair to certain groups of people.
- Chatbots that spread false information.
- Data breaches of sensitive information.

These incidents show that companies need to address the technological and ethical risks, Otherwise, they may face financial losses, damaged reputation, or loss of stakeholder trust.

Objectives

The main objective of this paper is to develop a framework that helps companies to manage the risks of using Artificial Intelligence. The objectives are both technical and ethical,

The specific objectives are:

- To identify and classify AI risks
- To propose a unified risk management framework
- To support corporate governance and decision-making
- To promote responsible and sustainable AI adoption

Contribution

The main contribution of this paper is the development of a unified framework that combines both technological and ethical perspectives of AI risk in corporate environments. Unlike previous approaches that focuses on a single aspect, this framework integrates both, showing how companies can use AI tools not only for productive improvements but also for risk detection and mitigation. Furthermore, it provides guidance for corporate boards on ethical oversight and promoting transparent and sustainable adoption of AI.

Methodology

The paper uses a conceptual and analytical approach rather than a dataset experiment. The methodology combines concepts from previous studies with new ideas to create a practical framework for managing AI risks in corporations.

First, the authors conduct a literature review to study how existing research addresses corporate risk management, AI ethics, and governance.

Second, they design a risk taxonomy, which classifies risks into two groups: technological risks and ethical risks. This classification becomes the foundation of the framework.

Third, the paper develops a unified framework that integrates these risks into a single model. The framework does not only highlight problems but also connects them with possible solutions.

Finally, the methodology includes case scenarios to show how the framework can work in real corporate settings.

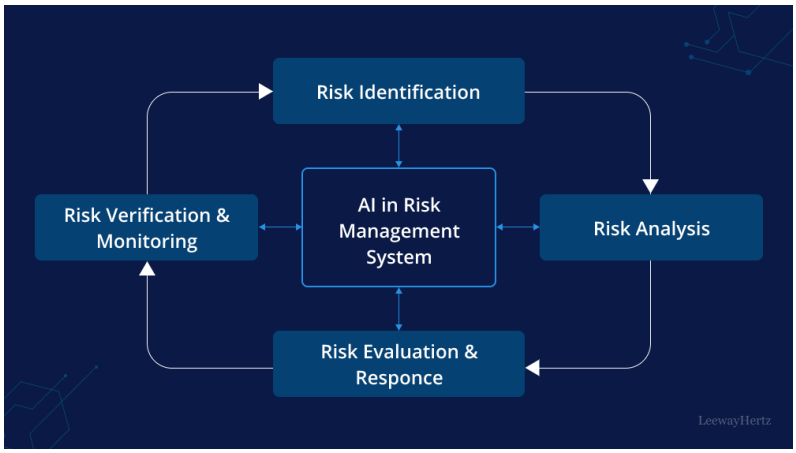

Figure 1: AI in Risk Management

Experimental Design

The proposed framework for managing technological and ethical risks in corporate AI applications is implemented through four core modules. Each module uses specific algorithms and mathematical models to detect risks, ensure ethical compliance, select the best strategies, and continuously improve performance over time.

- **Risk Identification Implementation**

  Random Forest algorithm is applied here. It is robust against overfitting and works well with large datasets having many variables.

  - Random Forest builds multiple decision trees and combines their outputs.
  - For classification, the final prediction is the mode of the classes; for regression, it is the mean prediction of individual trees.

  Mathematically, for an input feature vector X_i with risk indicator Y_i :

  $$Y_i = mode\ (\{f_1(X_i), f_2(X_i),..., f_n(X_i)\})$$

  Here f_j is the j^{th} decision tree, and n is the total number of trees. For regression, the mode is replaced with the mean.

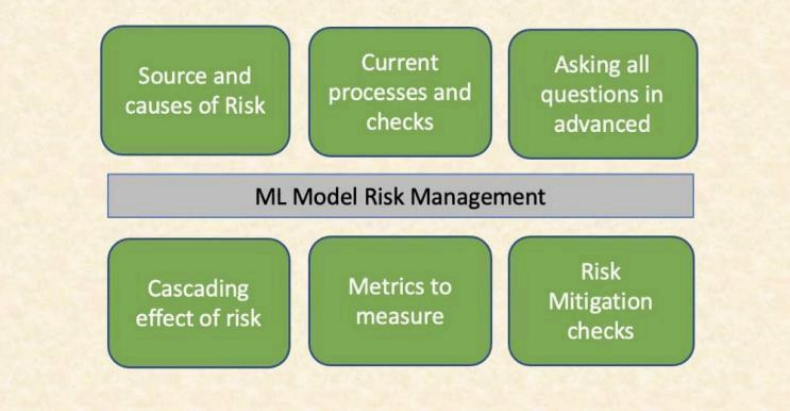

Figure 2: ML to Risk Management.

- **Ethical Oversight Mechanism**

  The framework introduces an **Ethical Auditing Tool** based on Natural Language Processing (NLP).

  - The tool checks AI decisions against ethical criteria.
  - Each criterion gets a score e_i, where $e_i = 1$ means fully compliant and $e_i = 0$ means not compliant.

  The ethical compliance score, S is then calculated as:

  $$S = \frac{1}{N}\sum_{i=1}^{N} e_i$$

  where N is the total number of ethical standards tested.

  The score tells whether the AI system meets required ethical guidelines.

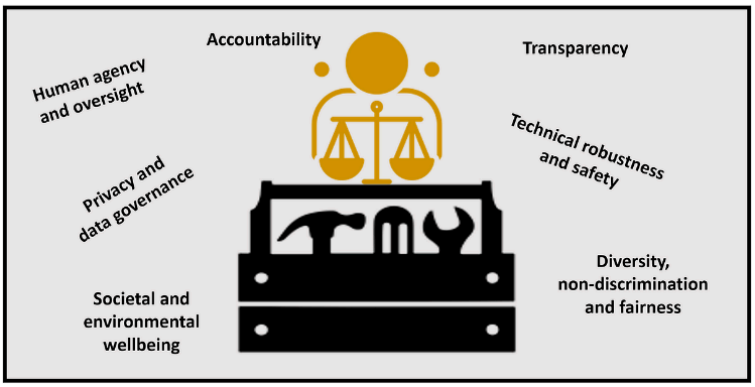

Figure 3: Ethical Oversight AI Tool.

- **Decision-making for Risk Mitigation**

  Once risks are detected, the framework uses decision trees and optimization algorithms to choose the best mitigation strategy. The goal is to maximize effectiveness while minimizing cost.

  The decision model is:

$$\max E(D) - \lambda_1 C(D) - \lambda_2 I(D)$$

  Here:

  - D = selected mitigation strategy
  - E(D) = effectiveness
  - C(D) = cost
  - I(D) = operational impact
  - λ_1, λ_2 = weights for cost and impact

  This approach ensures the chosen strategy balances risk reduction, cost, and business continuity.

Figure 4: Decision Making for Risk Mitigation.

- **Continuous Improvement Through Feedback**

  Finally, the framework uses reinforcement learning to continuously improve over time.

  - A reward function evaluates each mitigation effort based on its effectiveness, cost, and impact.
  - The model learns from outcomes to refine future decisions.

  The reward function is defined as:

$$R(O) = a - b.Cost(O) + c.Effectiveness(O) - d.I$$

  Here O represents the mitigation outcome, and a, b, c, d are parameters controlling the importance of cost, effectiveness, and impact.

  This way, strategies that deliver high effectiveness at low cost and impact get higher rewards.

Figure 5: Continuous Improvement Through Feedback.

Results

The proposed framework for AI and Corporate Risk Management was implemented using the four modules described earlier: Risk Identification, Ethical Oversight, Decision-making for Risk Mitigation, and Continuous Improvement through Feedback. The results show that the framework effectively detects technological and ethical risks, evaluates them, and suggests appropriate mitigation strategies while improving performance over time.

The following figure represents the evaluation of risk mitigation strategies,

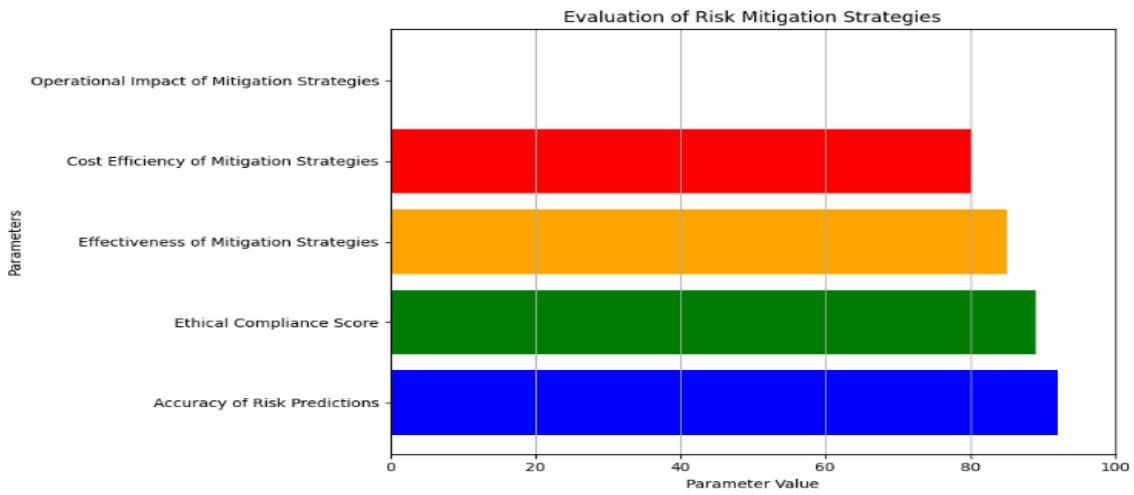

Figure 6: Performance Evaluation.

Findings:

- Accuracy of predictions and ethical compliance scored the highest, showing that the system is reliable and fair.
- Effectiveness and cost efficiency also performed decent.
- Operational impact of the strategies remained minimal.

Gap

While the proposed framework demonstrates strong performance in identifying risks, ensuring ethical compliance, and implementing cost-efficient mitigation strategies, several gaps and limitations persist:

- **Lack of Real-World Validation**

  The framework has been tested theoretically and through simulations, but real-world corporate cases across multiple industries are needed to validate its performance in practical conditions.

- **Dynamic Regulatory and Legal Issues**

  In real world scenario, data protection laws, AI ethics guidelines, and corporate governance policies differ significantly across countries and sectors.

- **Human and Organizational Factors**

  Issues such as organizational resistance to change, and cultural attitudes toward AI ethics are not deeply addressed.

Future Direction

To strengthen the proposed framework and make it more practical for real-world adoption, future work should focus on the following areas:

- **Real-World Case Studies**

  Deploy the framework in multiple industries and validate its performance under real world conditions.

- **Global Regulatory Adaptation**

  Extend the framework to account for different international data protection laws and AI ethics guidelines and sector-specific compliance rules. Build region-specific risk libraries so organizations can adapt the system to local laws automatically.

- **Integration with Corporate Ecosystems**

  Ensure minimal disruption and cost for organizations with existing infrastructure.

- **Long-Term Monitoring**

  Incorporate automated retraining pipelines for AI models as data patterns changes over time. Use predictive analytics to anticipate when performance degrades and take preventive measures.

- **Scalable Cloud-Based Deployment**

  Host the framework as a cloud service for scalability, security, and ease of updates, making it available to even small and medium enterprises.

References

[1] Naila Iqbal Qureshi, A. Garg, P. Singh, and N. Retzlaff, “AI and Corporate Risk Management: Identifying and Mitigating Technological and Ethical Risks,” vol. 169, pp. 1–5, Apr. 2024, doi: https://doi.org/10.1109/ickecs61492.2024.10617141.