PREDICTIVE ANALYTICS FFRESG RISK ASSESSMENT IN TEXTILE MANUFACTURING: A MULTIMODAL MACHINE LEARNING APPROACH

Sungho Kim[1], MD Rejoan Ahmed Rafi[2], Upamanyu Sengupta[3], Shahadat Hossain[2], Lorine Muenchow[3], MD Golam Sarwar[3], MD Abdul Kaium[4]

[1] Department of Computer Science, Korea University, Seoul, South Korea
[2] Department of Information Systems, Pacific States University, Los Angeles, USA
[3] Master of Business Administration, Pacific States University, Los Angeles, USA
[4] MS in Computer Science, Pacific States University, Los Angeles, USA

ABSTRACT

Environmental, Social, and Governance (ESG) risk assessment has become a critical requirement in global textile manufacturing due to increasing regulatory pressures, sustainability commitments, and brand accountability frameworks. Traditional ESG audits rely primarily on static inspections and retrospective reporting, limiting their ability to anticipate operational risks. This study proposes a multimodal machine learning framework for predictive ESG risk assessment by integrating heterogeneous data sources: manufacturer-submitted operational and compliance documents, worker happiness index surveys, and laboratory-generated water quality reports. Using supervised learning and multimodal feature fusion architectures, the model predicts ESG risk scores at the factory level across environmental and social dimensions. Experimental results demonstrate that multimodal models outperform unimodal baselines by 18–27% across F1-score and AUC metrics. The proposed framework enables early risk detection, supports proactive remediation, and offers scalable decision-support for brands, regulators, and sustainability stakeholders in textile supply chains.

KEYWORDS:

ESG analytics, textile manufacturing, multimodal machine learning, sustainability risk, worker wellbeing, water quality compliance

1. INTRODUCTION

The global textile and apparel industry is one of the most resource-intensive manufacturing sectors, accounting for significant water consumption, chemical discharge, labor intensity, and environmental externalities (Niinimäki et al., 2020). As sustainability expectations escalate, brands increasingly integrate Environmental, Social, and Governance (ESG) performance metrics into sourcing, compliance auditing, and supplier evaluation processes (Eccles et al., 2014). However, current ESG risk assessments remain largely reactive, relying on periodic inspections, document reviews, and historical compliance records. Such approaches lack predictive capability and fail to capture latent risks emerging between audit cycles.(Locke, 2013) Recent advances in predictive analytics and machine learning (ML) offer opportunities to transform ESG risk management from retrospective evaluation to proactive risk mitigation. In particular, multimodal machine learning, which integrates heterogeneous data sources such as numerical sensor data, textual reports, and human survey feedback, enables more comprehensive modeling of complex industrial systems. (Hastie et al., 2009) Despite growing interest in AI-enabled sustainability monitoring, the application of multimodal predictive analytics to ESG risk assessment in textile manufacturing remains underexplored. (Baltrušaitis et al., 2019) This study addresses this gap by proposing a multimodal ESG risk prediction framework that integrates three real-world data streams commonly available in textile manufacturing

DOI:10.5121/ijgca.2026.17111115

environments:

1. **Manufacturer-submitted operational documents** (e.g., production logs, compliance declarations, audit reports),
2. **Worker happiness index surveys**, capturing workforce wellbeing and social risk signals,
3. **Laboratory water quality test reports**, representing environmental compliance performance.

The primary research objective is to evaluate whether multimodal ML architectures improve ESG risk prediction accuracy compared to unimodal approaches and traditional scorecard-based assessments.

The contributions of this research are threefold:

1. Development of a domain-specific ESG risk modeling framework for textile manufacturing,
2. Empirical evaluation of multimodal fusion architectures using operational, social, and environmental data,
3. Demonstration of practical implications for predictive compliance, sustainability governance, and supply chain risk management.

2. Literature Review

2.1. ESG Risk Assessment In Textile Manufacturing

ESG frameworks such as the Global Reporting Initiative (GRI), Sustainability Accounting Standards Board (SASB), and UN Sustainable Development Goals (SDGs) provide standardized performance indicators for sustainability measurement (GRI, 2021; United Nations, 2015). In the textile sector, ESG risks are predominantly associated with excessive water consumption, chemical discharge, labor violations, unsafe working conditions, and governance failures in supplier operations. Studies indicate that conventional audit-based compliance mechanisms suffer from temporal gaps, information asymmetry, and incentive misalignment, leading to undetected violations and delayed remediation.

Moreover, ESG ratings in manufacturing often depend on static self-reported disclosures or third-party inspections conducted annually or semi-annually, which fail to reflect dynamic operational conditions. These limitations highlight the need for predictive and continuous ESG monitoring systems. (Locke et al., 2009)

2.2. Predictive Analytics In Manufacturing Systems

Predictive analytics has been extensively applied in manufacturing domains such as predictive maintenance, demand forecasting, energy optimization, and defect detection. Techniques including regression models, ensemble learning, deep neural networks, and time-series forecasting have demonstrated strong performance in anticipating operational failures and inefficiencies. However, most manufacturing analytics systems focus on productivity and quality optimization rather than sustainability or ESG outcomes. (Wuest et al., 2016) Recent studies suggest that predictive sustainability analytics can reduce emissions, optimize resource utilization, and improve occupational safety outcomes. Nevertheless, ESG risk modeling remains underdeveloped, particularly in labor-intensive and environmentally sensitive sectors such as textile manufacturing. (Wuest et al., 2016)

2.3. Multimodal Machine Learning

Multimodal machine learning integrates data from heterogeneous sources such as text, images, numerical records, and sensor streams. Fusion strategies are typically classified into:

- **Early fusion**, where raw features from multiple modalities are concatenated before learning,
- **Late fusion**, where independent models generate modality-specific predictions that are later aggregated,
- **Hybrid or attention-based fusion**, where dynamic weighting mechanisms learn cross-modal relationships.

Multimodal architectures have achieved state-of-the-art performance in healthcare diagnostics, autonomous systems, and financial risk modeling. However, their application to sustainability and ESG risk prediction in industrial manufacturing contexts remains limited, presenting an opportunity for methodological innovation. (Baltrušaitis et al., 2019)

2.4. Research Gap

Existing literature does not adequately integrate:

1. Worker wellbeing indicators with operational compliance data,
2. Laboratory environmental test results with social and governance metrics, 3. Multimodal predictive analytics for ESG risk modeling in textile manufacturing environments.
3. This study addresses these gaps by proposing and validating a multimodal ML-based ESG risk prediction framework grounded in real-world textile manufacturing data streams .

3. METHODOLOGY

3.1. Research Design

This study adopts a **quantitative, predictive modeling approach** using supervised machine learning. ESG risk prediction is framed as a classification and regression problem, where factories are assigned probabilistic ESG risk scores across environmental and social dimensions. The model is trained on historical compliance outcomes and audit findings labeled as low, medium, or high ESG risk.

3.2 Data Sources

Three complementary datasets were collected from participating textile manufacturing facilities operating in dyeing, finishing, and garment production units.

3.2.1. Manufacturer-Submitted Documents

Operational and compliance documents include:

- Monthly production and chemical usage logs,
- Occupational safety incident reports,
- Third-party compliance audit summaries,
- Environmental management system (EMS) disclosures.

These documents were primarily unstructured or semi-structured text data and numeric tabular records.

3.2.2. Worker Happiness Index Surveys

Worker well-being data were collected through standardized anonymous surveys administered quarterly. The survey instrument measured:

- Job satisfaction,
- Perceived workplace safety,
- Supervisor support,
- Wage fairness,
- Psychological well-being.

Responses were aggregated into a normalized **Happiness Index Score (0–100)** per factory per quarter.

3.2.3. Laboratory Water Quality Reports

Certified laboratories conducted monthly effluent testing on wastewater discharge samples. Measured indicators included:

● Chemical Oxygen Demand (COD),
● Biological Oxygen Demand (BOD),
● pH,
● Total Dissolved Solids (TDS),
● Heavy metals concentration.

These indicators were benchmarked against national environmental standards to generate compliance labels.

3.3 Data Preprocessing

- **Textual documents** were processed using tokenization, stop-word removal, and
- Transformer-based embeddings.
- **Survey data** were normalized and temporally aligned with production cycles.
- **Water quality data** were standardized and transformed into rolling averages and violation flags.
- Missing values were imputed using median imputation for numerical variables and masked embeddings for textual features.

All datasets were synchronized at the factory-month granularity.

3.4 Feature Engineering

Key engineered features included:

- Frequency of compliance violations per quarter,
- Change rate in worker happiness index,
- Effluent exceedance ratios,

Textual sentiment polarity and topic embeddings from audit narratives. These features were mapped to ESG pillars as follows:

Esg Pillar Feature Examples

Environmental COD exceedance, water consumption per unit, chemical intensity
Social Worker happiness index, incident reports, and absenteeism
Governance Audit frequency, corrective action closure rate

3.5 Model Architecture

A **multimodal deep learning architecture** was implemented, consisting of:

- **Text Encoder:** Transformer-based embeddings for manufacturer documents,
- **Time-Series Encoder:** LSTM networks for water quality indicators,
- **Numerical Encoder:** Dense neural layers for worker happiness index and tabular metrics.

These modality-specific embeddings were fused using an **attention-based fusion layer**, followed by fully connected layers for ESG risk classification and regression output.

Baseline unimodal models were trained separately on each modality using random forest classifiers and gradient boosting regressors.

3.6 Evaluation Metrics

Models were evaluated using:

- Accuracy,
- Precision,
- Recall,
- F1-score,

Area Under the Receiver Operating Characteristic Curve (AUC),
Mean Absolute Error (MAE) for continuous risk score prediction.

Cross-validation with an 80:20 train-test split was used, ensuring temporal separation between training and test datasets.

4. EXPERIMENTAL RESULTS

4.1. Predictive Performance

Table 1 summarizes model performance across ESG risk classification tasks. **Table1. Model Performance Comparison**

Model Type F1-Scor e AUC MAE
Water Quality Only 0.71 0.76 0.29
Worker Happiness Only 0.68 0.72 0.31
Manufacturer Documents Only 0.74 0.79 0.27

Multimodal Fusion Model **0.88 0.91 0.16**
The multimodal fusion model outperformed unimodal baselines by:

- 18–27% improvement in F1-score,
- 15–19% improvement in AUC,
- 40–45% reduction in prediction error.

4.2 Feature Contribution Analysis

SHAP (SHapley Additive exPlanations) analysis revealed that:

- Sudden drops in the worker happiness index strongly predicted social compliance violations,
- COD and BOD exceedance patterns were the strongest predictors of environmental risk,
- Negative sentiment in audit narratives correlated with governance non-compliance escalation.

4.3 Early Warning Capability

The model successfully identified 72% of high-risk ESG events at least one audit cycle prior to formal compliance failure, demonstrating its potential as a real-time early warning system for sustainability governance.

5. DISCUSSION

The results confirm that multimodal data integration significantly enhances ESG risk prediction accuracy in textile manufacturing contexts. While traditional compliance approaches treat environmental, social, and governance indicators as siloed domains, this study demonstrates the existence of meaningful cross-domain relationships. For instance, declining worker wellbeing scores were often observed to precede environmental violations, indicating broader systemic governance issues within factory management structures. From a sustainability governance perspective, the proposed framework enables proactive risk mitigation through early detection, supports targeted intervention planning by identifying dominant risk drivers, and facilitates dynamic ESG reporting aligned with regulatory and buyer expectations. Furthermore, the inclusion of worker happiness index data introduces a human-centered dimension to ESG analytics, extending beyond physical safety indicators to encompass psychological and organizational wellbeing, an aspect that has been underrepresented in prior ESG modeling research.

6. ETHICAL CONSIDERATIONS

The study adheres to established ethical data governance principles by ensuring that all worker surveys were anonymized and aggregated to protect individual privacy. Manufacturer data were processed under strict confidentiality agreements to safeguard commercially sensitive information. In addition, the models were evaluated for potential bias amplification across different factory sizes and geographic regions to support fair and responsible use. Explainable AI techniques were applied to enhance model interpretability, enabling compliance managers to better understand, trust, and appropriately act on the model outputs.

7. LIMITATIONS

Several limitations should be acknowledged. Data availability and reporting accuracy varied across manufacturers, which may have affected the consistency and reliability of the inputs used for model training. In addition, the study relied on supervised labels derived from audit outcomes, which can reflect institutional bias and may not fully capture underlying ESG risks. Real-time sensor integration was not included in the current framework, limiting the model's ability to detect immediate or rapidly evolving compliance issues. Furthermore, external geopolitical and regulatory risks were not explicitly modeled, despite their potential impact on ESG performance. Future research may address these limitations by integrating IoT-based emissions sensors, blockchain-enabled traceability data, and enhanced supplier-tier visibility to improve predictive accuracy and robustness.

8. CONCLUSION

This study demonstrates that **multimodal machine learning significantly improves ESG risk assessment accuracy in textile manufacturing environments** by integrating manufacturer-submitted operational documents, worker happiness index surveys, and laboratory water quality reports. The proposed framework enables predictive sustainability governance, facilitating earlier intervention, improved compliance performance, and enhanced social and environmental accountability.

By transforming ESG monitoring from reactive audits to proactive risk prediction, this research contributes to the advancement of data-driven sustainability management in global textile supply chains and establishes a scalable foundation for intelligent ESG governance systems.

9. FUTURE RESEARCH DIRECTIONS

Future work may explore the integration of real-time IoT-enabled emissions monitoring to enable continuous and fine-grained environmental risk detection. In addition, causal inference models could be developed to better understand and quantify the relationships between operational decisions and resulting ESG outcomes, moving beyond correlation-based predictions.

Reinforcement learning approaches also present promising opportunities for designing adaptive sustainability intervention strategies that dynamically optimize compliance and resource efficiency over time. Finally, extending the framework to include cross-country regulatory benchmarking and supply-chain tier propagation modeling would allow for a more comprehensive assessment of ESG risks across global, multi-tier textile supply networks.

REFERENCES

[1] Allwood, J. M., Laursen, S. E., De Rodríguez, C. M., &Bocken, N. M. P. (2006). Well dressed? The present and future sustainability of clothing and textiles in the United Kingdom. *Journal of the Textile Institute, 97*(3), 1–18.
[2] Baltrušaitis, T., Ahuja, C., & Morency, L. P. (2019). Multimodal machine learning: A survey and taxonomy. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 41*(2), 423–443.
[3] Eccles, R. G., Ioannou, I., & Serafeim, G. (2014). The impact of corporate sustainability on organizational processes and performance. *Management Science, 60*(11), 2835–2857.
[4] GRI. (2021). *GRI sustainability reporting standards*. Global Reporting Initiative.
[5] Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning* (2nd ed.). Springer.
[6] Locke, R. M. (2013). *The promise and limits of private power*. Cambridge University Press.

[7] Locke, R. M., Amengual, M., & Mangla, A. (2009). Virtue out of necessity? Compliance, commitment, and the improvement of labor conditions in global supply chains. *Politics & Society, 37*(3), 319–351.
[8] Niinimäki, K., Peters, G., Dahlbo, H., Perry, P., Rissanen, T., & Gwilt, A. (2020). The environmental price of fast fashion. *Nature Reviews Earth & Environment*, *1*(4), 189–200.
[9] United Nations. (2015). *Transforming our world: The 2030 agenda for sustainable development*.
[10] Wuest, T., Weimer, D., Irgens, C., & Thoben, K. D. (2016). Machine learning in manufacturing: Advantages, challenges, and applications. *Production & Manufacturing Research, 4*(1), 23–45.