

From Data to Strategy: Machine Learning for Identifying Risk Patterns in Amazon's Operations

G. Gardenal, C. Gianluca - Confirmed, L. Barattini

Amazon EU SARL | 38 AVENUE JOHN F. KENNEDY, L-1855 LUXEMBOURG

Abstract—This one-pager presents a proposal for a bachelor's thesis aimed at applying machine learning techniques to analyze and predict incidents inside Amazon's operational framework. The scope of this document is also to involve stakeholders from both the academic and corporate field, highlighting the potential advantages and strategic insights that such a research could provide.

keywords—Risk Management, Machine Learning, Operational Risk

Table of Contents - To Be Confirmed

Abstract

Introduction

Chapter 1: Foundations of Operational Risk Management

- Definitions
- Defining our focus area
- Amazon's Focus: Key Risk Types
- Categorizing Risks: SEV 1-5 Framework

Chapter 2: Amazon's Risk Management Strategies

- Approaches to Risk Management at Amazon
- Advantages of a Data-Driven Risk Management Approach

Chapter 3: Research Framework and Methodology

- Key Research Questions
- Metrics and Performance Thresholds for Model Evaluation

Chapter 4: Technical Implementation and Model Development

- Data Cleaning Techniques and Challenges
- Data Preparation and Standardization
- Feature Engineering
- Documentation of Model Selection
- Iterative Model Development and logging
- Defining and Validating Model's KPIs
- Setting and Testing Performance Thresholds

Chapter 5: Discussions and Conclusions

- Analysis of Findings and Identified Patterns
- Lessons Learned and Implications for Future Research

Appendix

References

1. Context

As an intern within Amazon's CF FP&A variable reporting team, I immersed myself into comprehensive planning and reporting cycles aggregating financial inputs across multiple countries. Recently, my attention shifted towards the SEV1&2 cost bucket, crucially shaped by EMEA's operational risk management prevention policies. Guided by the possibility of enhancing these strategies, I am keen to investigate recurring patterns within this area and facilitate a targeted action plan in collaboration with the EMEA S&LP Team.

2. Action Plan - Definitive

1. **Literature Review and Precedent Analysis:** Perform a comprehensive analysis of internal Risk Management documents, including established policies, standard procedures, cost allocation practices, and relevant definitions. ✓

2. **Data Acquisition and Preliminary Analysis:** Partner with the EU S&LP team to obtain and scrutinize historical data. ✓

3. Model Implementation and coding

- 3.1. Data Cleaning ✓
- 3.2. Data Preparation ✓
- 3.3. Feature Selection
- 3.4. Data Quality and Accuracy
- 3.5. Data Privacy Requirements
- 3.6. Model Selection Documentation
- 3.7. Model Development Iterations Logbook
- 3.8. Define Model's KPI
- 3.9. Define Acceptable Performance Threshold
- 3.10. Confirm and Finalize Model Results

4. **Chapter Development and Refinement** Compile research findings into structured chapters to systematically present research methodologies, data analysis and findings.

3. Thesis Abstract - Definitive

In this thesis, machine learning techniques are applied to explore the aggregation of incidents within Amazon's operational framework, aiming to identify and analyze patterns that could inform risk management strategies. Through a comprehensive analysis, the research highlights specific conditions under which these events are more likely to occur. The findings propose several focus areas for Amazon's risk management team to potentially reduce both the frequency and impact of these occurrences, emphasizing data-driven decision-making.

4. Access and Visibility

As stipulated by Ca' Foscari University of Venice[3] Comma IV.I, this thesis will be subject to an embargo period of 36 months following its submission. During this time, the document will not be publicly accessible. This measure ensures the confidentiality of proprietary data.

5. Support from Current Literature

Recent studies underscore the significant potential of leveraging data analytics in operational risk management. Scandizzo's exploration of risk mapping and indicators, Araz's detailed discussion on data analytics for risk management, and Samad-Khan's insights into modern risk management practices all affirm the transformative impact of data-driven approaches in this field [1], [2], [4].

References

- [1] S. Scandizzo, *Risk Mapping and Key Risk Indicators in Operational Risk Management*. (English). 2005, vol. 51, pp. 230–256. DOI: <https://doi.org/10.1111/j.0391-5026.2005.00150.x>.
- [2] O. M. Araz, "Data Analytics for Operational Risk Management. (English)", *Decision Sciences*, vol. 51, no. 6, 2020. DOI: <http://dx.doi.org/10.1111/deci.12443>.
- [3] "Ca' foscari university of venice". (), [Online]. Available: <https://www.unive.it/pag/31196/>.
- [4] A. Samad-Khan. "Modern operational risk management". (), [Online]. Available: <https://www.sbp.org.pk/bsrvd/pdf/KnowledgeSharing/Reading%20Material%20for%20workshop%20on%20ORM%20May%2018-22,%202009/General%20Articles/Modern%20ORM.pdf>.