



An Integrated Business Intelligence Framework for Sustainable Risk Mitigation

Harith Yas

Faculty of Management, Universiti Teknologi Malaysia Johor Bahru, Johor, Malaysia

Email: harith.albayati@yahoo.com

Abstract

This research paper examines the critical juncture of business intelligence and sustainable risk management in response to the increasing challenges faced by modern businesses. Our study recognizes that organizations must navigate uncertainties while prioritizing sustainability. It focuses on analyzing credit risk data. We present a comprehensive examination of predictive performance using Logistic Regression, Decision Tree, and K-Nearest Neighbors classifiers augmented by the Synthetic Minority Over-sampling Technique (SMOTE) for class rebalancing. The empirical findings presented through detailed tables and figures reveal intricate relationships and patterns within the data. This research also contributes to the broader discourse on responsible business practices by highlighting the integration of business intelligence in sustainable risk mitigation. Moreover, comparative analysis of machine learning algorithms under various resampling techniques further strengthens the framework's reliability.

Keywords: Business Intelligence; Risk Management, Sustainability; Risk Governance; Resilient Enterprises; Social Responsibility; Economic Sustainability.

1. Introduction

In the last few years, the intersection of business intelligence and sustainability has become a focal point for organizations aiming to navigate the complex landscape of contemporary risks. Enterprises are compelled to adopt a more proactive and integrated approach to risk management as global markets evolve and stakeholder expectations shift. Traditional risk mitigation strategies are often inadequate in addressing the multifaceted challenges posed by environmental, social, and economic factors. The integration of business intelligence (BI) tools has gained prominence as a strategic imperative for organizations seeking to harness the power of data-driven decision-making [1-3]. However, the application of BI in the context of sustainable risk management remains an underexplored domain. By amalgamating BI methodologies with sustainability principles, this research endeavors to bridge this gap, presenting a novel framework that not only identifies potential risks but also empowers businesses to proactively mitigate them. Such an approach aligns with the evolving paradigm of corporate responsibility, where a proactive stance on sustainability is integral to long-term success [4].

The need for a coherent framework that considers the interplay between environmental, social, and economic factors becomes increasingly apparent as businesses grapple with an array of interconnected risks. The proposed integrated BI framework is designed to go beyond traditional risk silos, fostering a holistic understanding of the dynamic landscape in which modern enterprises operate. This research aims to contribute to a paradigm shift in risk management paradigms by acknowledging the intricate relationships between sustainability and risk, offering a comprehensive solution that not only safeguards against potential threats but also positions organizations as responsible contributors to societal well-being [5].

This research holds relevance within the context of the Business Sustainability journal, as it aligns with the journal's commitment to advancing knowledge and understanding of sustainable business practices. By elucidating the interconnectivity between business intelligence and sustainable risk management, this paper seeks to provide a valuable contribution to the ongoing discourse on responsible and resilient corporate governance.

2. Methodology

This section is the most important part of the study as it gives a clear roadmap for designing, executing and analyzing our research. Logistic Regression is used to initiate our analysis because it is a foundational algorithm that models the probability of binary outcomes. Logistic Regression is based on regression analysis principles and uses logistic function to map predicted values into probabilities, which helps in understanding the likelihood of credit default in our sustainable risk management context. This paragraph explains why Logistic Regression is suitable for our predictive modeling objectives [6].

The next section introduces the Decision Tree algorithm as a powerful tool for classification tasks with interpretability and simplicity in mapping out decision rules. Decision Trees are based on recursive partitioning principles that split the dataset iteratively based on features, creating a tree-like structure that helps to understand the decision-making process. This paragraph highlights the main principles of Decision Trees and their importance in identifying patterns in credit risk data for sustainable risk detection [7-8].

The methodology follows Decision Tree and incorporates the K-Nearest Neighbors (KNN) Classifier, which is a non-parametric algorithm based on proximity-based classification. KNN classifies a data point by considering the classes of its k-nearest neighbors, making it particularly effective for patterns where proximity is indicative of similar outcomes. This paragraph explains the theoretical basis of KNN, emphasizing its suitability to the complexities of sustainable risk management in our credit risk dataset [9-11].

To address potential imbalances in the class distribution within the credit risk dataset and enhance the robustness of our predictive models, the Synthetic Minority Over-sampling Technique (SMOTE) is systematically applied. As an oversampling technique, SMOTE synthesizes artificial instances of the minority class by interpolating between existing instances, effectively mitigating the class imbalance challenge. This paragraph explains SMOTE's main principles and highlights its role in rebalancing the dataset, thus contributing to better performance and generalization capabilities of Logistic Regression, Decision Tree, and K-Nearest Neighbors classifiers in sustainable detection of business risks.

3. Results and Discussion

This section is not just about presenting raw data, but it is a melting pot where empirical evidence meets theoretical foundations to create a nuanced discussion that explores the implications, limitations, and future research directions. The dataset used in this study is a publicly available case study on credit risk with comprehensive details about credit borrowers. The dataset consists of key variables that give a complete picture of individuals' credit profiles such as age, education level, tenure with the current employer, number of years at the same address, customer income, debt-to-income ratio, credit-to-debt ratio, other debts and a binary indicator showing whether the customer has defaulted in the past (where 1 means default while 0 means no default history). This dataset provides a strong basis for our empirical analysis enabling us to examine the complex dynamics of credit risk through our integrated business intelligence framework for sustainable risk mitigation. The results of our statistical analyses are systematically presented and visually summarized in Table 1. This tabular representation serves as a concise and structured means to communicate the key outcomes of our empirical investigation. Each column and row within the table corresponds to a specific variable or statistical measure, offering a clear and organized snapshot of the relationships and patterns identified [14-17]. With the employment of visual elements such as tables, we aim to enhance the accessibility and interpretability of our results, facilitating a comprehensive understanding of the nuanced statistical insights derived from the application of our integrated business intelligence framework to the public case study on credit risk data [18].

Table 1: Comprehensive Statistical Analysis of Credit Risk Data Using the Integrated Business Intelligence Framework

	age	ed	employ	address	income	debtinc	creddebt	othdebt	default
count	850	850	850	850	850	850	850	850	700
mean	35.02941	1.710588	8.565882	8.371765	46.67529	10.17165	1.576805	3.078789	0.261429
std	8.041432	0.927784	6.777884	6.895016	38.54305	6.719441	2.12584	3.398803	0.439727
min	20	1	0	0	13	0.1	0.011696	0.045584	0
25%	29	1	3	3	24	5.1	0.382176	1.045942	0
50%	34	1	7	7	35	8.7	0.885091	2.003243	0
75%	41	2	13	12	55.75	13.8	1.89844	3.903001	1
max	56	5	33	34	446	41.3	20.56131	35.1975	1

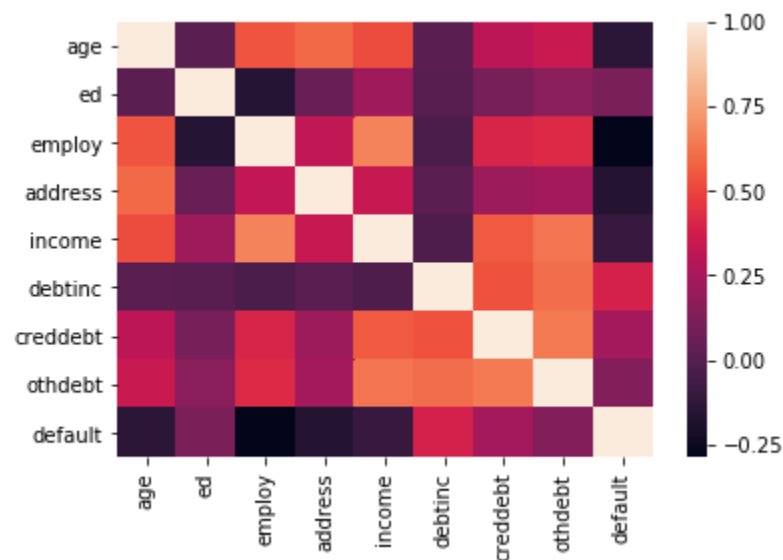


Figure 1: Correlation Analysis for Key Variables in Credit Risk Data

In Figure 1, we delve into a detailed distribution analysis that provides a visual representation of crucial patterns within the credit risk data. This graphical exploration serves as a valuable complement to the quantitative insights presented in the statistical analyses. By examining the distribution of key variables, such as age, education level, tenure with the current employer, and others, we gain a nuanced understanding of the inherent patterns and variations within the dataset. The visualizations in Figure 1 not only enhance the interpretability of our findings but also offer a holistic perspective on the distributional characteristics of the variables under scrutiny. In Figure 2, we present comprehensive correlation maps that offer a visual representation of the interrelationships among key variables within the credit risk dataset. These correlation maps serve as a powerful analytical tool, shedding light on the strength and directionality of associations between variables. By employing a color-coded scheme, the maps provide an intuitive depiction of correlation coefficients, facilitating the identification of patterns and dependencies. Through these visualizations, we aim to uncover the intricate connections that exist within the dataset, offering valuable insights into how factors such as age, education level, tenure with the current employer, and others correlate with each other.

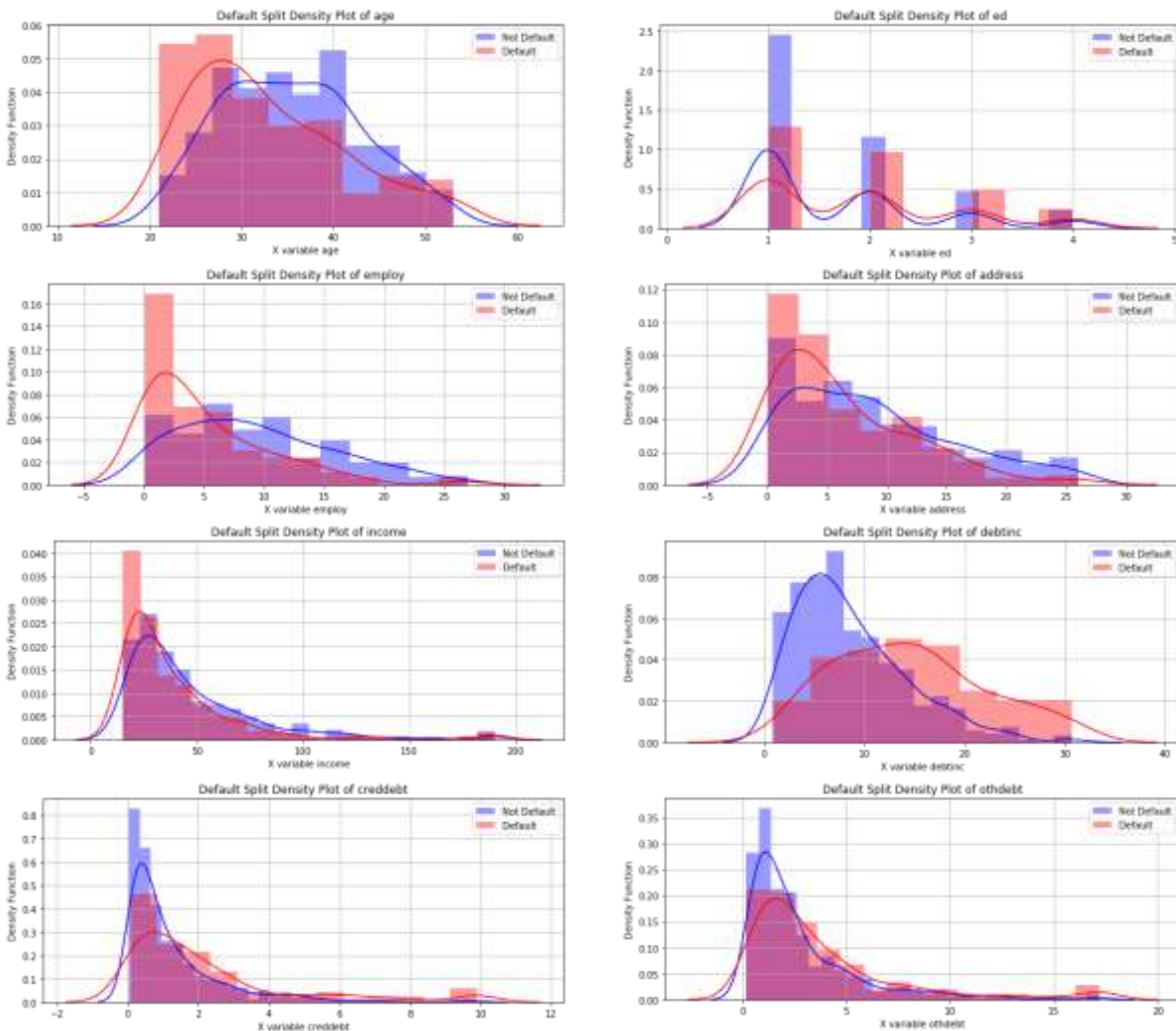


Figure 2: Distribution Analysis for Key Variables in Credit Risk Data

Table 2 presents a comprehensive overview of the prediction performance of various machine learning algorithms under different resampling techniques, namely under-sampling, oversampling, and Synthetic Minority Over-sampling Technique (SMOTE). The application of these resampling methods is pivotal in addressing class imbalances within the credit risk dataset, and our table meticulously captures the nuanced outcomes of this process. By systematically organizing and juxtaposing the performance metrics of machine learning algorithms—ranging from accuracy and precision to recall and F1-score—across the diverse resampling strategies, we provide a detailed analysis of how these techniques impact the predictive capabilities of the models. This comparative assessment not only underscores the robustness of our integrated business intelligence framework but also contributes to the broader discourse on the effectiveness of resampling methodologies in enhancing the predictive accuracy and reliability of machine learning algorithms in the domain of sustainable risk management.

Table 2: Prediction Performance of Machine Learning Algorithms under Various Resampling Techniques

method	Undersampling	Oversampling	SMOTE
Logistic Regression	0.697996	0.741251	0.752224
Decision Tree	0.597051	0.907552	0.806122
KNN Classifier	0.648139	0.803186	0.830017

4. Conclusion

This research endeavors to provide a comprehensive framework for sustainable risk management through the integration of advanced business intelligence methodologies. The exploration of Logistic Regression, Decision Tree, and K-Nearest Neighbors classifiers, coupled with the strategic application of the Synthetic Minority Over-sampling Technique (SMOTE) for class rebalancing, offers valuable insights into the intricate dynamics of credit risk within a sustainability context. Our empirical findings, systematically presented in Tables and Figures, illuminate the nuanced relationships and patterns identified through statistical analyses, distribution examinations, and correlation mapping. By bridging the realms of business intelligence and sustainable risk management, this study not only contributes to the scholarly discourse but also provides practical implications for organizations seeking resilient risk mitigation strategies.

References

- [1] Marar, Fatima Murshed Al, and Haitham Nobanee. n.d. "Business Intelligence, Sustainability and Risk Management."
- [2] Žigienė, Gerda, Egidijus Rybakovas, and Robertas Alzbutas. 2019. "Artificial Intelligence Based Commercial Risk Management Framework for SMEs." *Sustainability* 11 (16): 4501.
- [3] Goede, Roelien. 2021. "Sustainable Business Intelligence Systems: Modelling for the Future." *Systems Research and Behavioral Science* 38 (5): 685–95.
- [4] Oduoza, C F. 2020. "Framework for Sustainable Risk Management in the Manufacturing Sector." *Procedia Manufacturing* 51: 1290–97.
- [5] Bolos, Marcel Ioan, Diana Claudia SABAU-POPA, Ioana-Alexandra BRADEA, Camelia DELCEA, and Emil SCARLAT. 2016. "A BUSINESS INTELLIGENCE INSTRUMENT FOR DETECTION AND MITIGATION OF RISKS RELATED TO PROJECTS FINANCED FROM STRUCTURAL FUNDS." *Economic Computation & Economic Cybernetics Studies & Research* 50 (2).
- [6] Ullah, Fahim, Siddra Qayyum, Muhammad Jamaluddin Thaheem, Fadi Al-Turjman, and Samad M E Sepasgozar. 2021. "Risk Management in Sustainable Smart Cities Governance: A TOE Framework." *Technological Forecasting and Social Change* 167: 120743.
- [7] Seddigh, Mohammad Reza, Sajjad Shokouhyar, and Fatemeh Loghmani. 2023. "Approaching towards Sustainable Supply Chain under the Spotlight of Business Intelligence." *Annals of Operations Research* 324 (1–2): 937–70.
- [8] Serrano, Frederico José Pereira. 2023. "Risk Management In Supply Chain Integration Using A Business Intelligence Optimization Approach."
- [9] Piekiet Weeserik, Bram, and Marco Spruit. 2018. "Improving Operational Risk Management Using Business Performance Management Technologies." *Sustainability* 10 (3): 640.
- [10] Chang, Victor, Raul Valverde, Muthu Ramachandran, and Chung-Sheng Li. 2020. "Toward Business Integrity Modeling and Analysis Framework for Risk Measurement and Analysis." *Applied Sciences* 10 (9): 3145.
- [11] Kabra, Gaurav, Vinit Ghosh, and A Ramesh. 2018. "Enterprise Integrated Business Process Management and Business Intelligence Framework for Business Process Sustainability." In *Green Initiatives for Business Sustainability and Value Creation*, 228–38. IGI Global.
- [12] Zheng, Jiehui, and Haliyana Khalid. 2022. "The Adoption of Enterprise Resource Planning and Business Intelligence Systems in Small and Medium Enterprises: A Conceptual Framework." *Mathematical Problems in Engineering* 2022.
- [13] Charles, Vincent, Pratibha Garg, Neha Gupta, and Mohini Agarwal. 2023. *Data Analytics and Business Intelligence: Computational Frameworks, Practices, and Applications*. CRC Press.
- [14] Badwan, Nemer, Lamees Al-Zoubi, and Suhaib Al-Khazaleh. 2023. "Financial Risk Management in the Supply Chain Using Business Intelligence and Big Data." *International Journal of Management and Business Research* 7 (1): 1–27.
- [15] Avendaño, E, A Moreno, G Sanchez, and M Villamil. 2020. "Business Intelligence and Visual Analytics-Based Strategy for Disaster Risk Observatory Development." *Prospectiva (1692-8261)* 18 (1).

- [16] Spedding, Linda S, and Adam Rose. 2007. *Business Risk Management Handbook: A Sustainable Approach*. elsevier.
- [17] Charneira, Nuno Miguel Valentim. 2022. “Monitoring Smart Sustainable Cities: Business Intelligence Approach to Smart Sustainable Cities.”
- [18] Torres, Germán Sánchez, Mar\`ia del Pilar Villamil, Andres Moreno, and Eduard Avendaño. 2020. “Business Intelligence and Visual Analytics-Based Strategy for Disaster Risk Observatory Development.” *Prospectiva* 18 (1): 2.