CHEMICAL ENGINEERING TRANSACTIONS

VOL. 106, 2023

A publication of AIDIC
The Italian Association of Chemical Engineering
Online at www.cetjournal.it

Guest Editors: Jeng Shiun Lim, Nor Alafiza Yunus, Peck Loo Kiew, Hon Huin Chin

ISBN 979-12-81206-05-2; **ISSN** 2283-9216

DOI: 10.3303/CET23106002

A Rule-based Model for Predicting Airline Financial Performance from Environmental, Social, and Governance Data

Josiah Kurt B. Aviso[a], Jonna C. Baquillas[b], Kathleen B. Aviso[c,*], Tsai-Chi Kuo[d], Hsiao-Min Chen[e], Raymond R. Tan[c]

[a]Department of Software Technology, College of Computer Studies, De La Salle University, Manila, Philippines
[b]Department of Marketing and Advertising, De La Salle University, Manila, Philippines
[c]Department of Chemical Engineering, De La Salle University, Manila, Philippines
[d]Department of Industrial Management, National Taiwan University of Science and Technology, No.43, Sec.4, Keelung Rd., Taipei, 106335, Taiwan
[e]Department of Marketing and Logistics Management, Chihlee University of Technology, Taipei, 220305, Taiwan
kathleen.aviso@dlsu.edu.ph

Environment, social, and governance (ESG) considerations have become a necessity for businesses. A company's environmental and community initiatives have been found to greatly influence customer perception. This is even more critical for industries that are difficult to decarbonize like the aviation industry. There has been little investigation on the role of ESG strategies on company financial performance which can dictate the sustainability of initiative implementation. This work uses ESG performance indicators to develop a rule-based model for predicting company financial performance as measured by return on assets (ROA). Results suggest that the most critical attributes of the ESG framework are Innovation, Workforce, Human Rights, Product Responsibility, Shareholders, and the aggregate ESG score. The best–performing model correctly predicts 15 out of 28 of the validation data (53.57 %). A rule of interest is that which states IF (Human Rights = Average) THEN (ROA = Average). It had the highest coverage for both training and validation data with a certainty of 61 %, and a prediction accuracy of 71.4 %, highlighting the importance of Human Rights on firm value.

1. Introduction

The aviation industry is a major contributor to climate change (Fahey et al., 2016) and decarbonizing the sector remains a challenge (Amankwah-Amoah et al., 2022). It accounts for about 10 % of the global greenhouse gas (GHG) emissions for domestic and international transport (Bergero et al., 2023), consisting of cargo and passenger movements (IPCC, 2022). Despite the environmental impacts, the aviation industry is vital in tourism and cargo transport (Fahey, 2016). Tourism is seen as an important economic driver and has a significant role in poverty alleviation (Perryman et al., 2022). Air cargo is essential in the movement of perishable goods, high value items, and production operation support, a critical role in the global economy (Gupta, 2017).

Environment, Social, and Governance (ESG) performance has received notable global attention in recent years (Bang et al., 2023), with discussions often crossing paths with Corporate Social Responsibility (CSR) (Gillan et al., 2021), which focuses on organizational commitment to pursue strategies that support society while doing good business (Tsang et al., 2022). Despite prominence gained in recent years, no standard meaning has been established to differentiate these two in academic research, with most using them interchangeably (Torres et al., 2023). CSR prioritizes better corporate citizenship with indirect reference to governance linked with environmental and social strategies, while ESG explicitly includes governance issues, thus being more expansive (Gillan, 2021).

A company's ESG details corporate sustainability initiatives' impact in response to the demand of investors and stakeholders (Tsang et al., 2022). In the airline industry, Kuo et al. (2021) showed that ESG performance has a significant relationship on Return on Assets (ROA), moderated by airline ownership type. The financial

Paper Received: 14 June 2023; Revised: 16 July 2023; Accepted: 31 July 2023
Please cite this article as: Aviso J.K.B., Baquillas J.C., Aviso K.B., Kuo T.-C., Chen H.-M., Tan R.R., 2023, A Rule-based Model for Predicting Airline Financial Performance from Environmental, Social, and Governance Data, Chemical Engineering Transactions, 106, 7-12 DOI:10.3303/CET23106002

performance of full-service air carriers is enhanced by environmental and social initiatives, while low-cost carriers are enhanced by increased firm size and environmental initiatives (Yang and Baasandorj, 2017). Few studies have investigated sustainability reporting in the airline industry (Zieba and Johansson 2022), and fewer still in understanding the effect of ESG scores on the financial performance of airline companies.

Recently, artificial intelligence (AI) and machine learning (ML) techniques have been used for various sustainability applications such as sustainable business management (Khan et al., 2022) and detecting patterns for energy use and GHG emissions in cities (Aviso et al., 2021). The use of interpretable ML models such as rough set–based ML (RSML), in contrast to black box models, is recommended as it provides better decision support to decision makers (Rudin, 2019). A search in the Scopus database using the keywords TITLE-ABS-KEY (predicting AND financial AND performance AND from AND esg) yields only 8 published documents suggesting that more work is needed in the area. This work explores the predictive power of disaggregated components of ESG instruments on financial performance. Section 2 formally defines the problem addressed in this work, Section 3 provides a brief overview of rough set machine learning, while Section 4 describes the case study examined. The results and discussion are then given in Section 5 while conclusions and directions for future work are discussed in Section 6.

2. Problem Statement

The data set is assumed to be heterogeneous and may consist of disjoint or overlapping rough sets (clusters). Discoverable patterns within each rough set can be approximated by rules, but it is possible that some examples in the data set do not fit any of these patterns. The problem can be stated formally as follows:

- Given a set of condition attributes according to the Thomson Reuter's ESG performance framework;
- Given the decision attribute Return on Assets (ROA) to represent a company's financial performance;
- Given the performance of airline companies in the condition and decision attributes;

The problem is to determine a rule-based model that can predict the ROA of an airline company.

3. Rough Set–Based Machine Learning

Rough set theory (RST) was developed by Pawlak (1982) to lay down the mathematical foundations for the approximate equality of sets, which can be useful in some branches of AI. It defines the approximation space, A=(U,R) where U refers to the universal set and R is an equivalence relation. If objects $(x,y) \in U$ and $(x,y) \in R$ then x and y are said to be indiscernible in A. This indiscernibility relation has been proven useful for AI applications like classification and clustering. For practical applications, objects defined in the approximation space A, can be organized into an information system. When R is further partitioned into condition, C, and decision, D, attributes, then the information system is transformed into a decision table where decision attributes can be used for creating subsets or decision classes. Objects indiscernible in A can be classified into subset or decision class X by defining the lower $\underline{Apr_A}(X)$ and upper $\overline{Apr_A}(X)$ approximation of X. $\underline{Apr_A}(X)$ contains all objects that definitely belong in X while $\overline{Apr_A}(X)$ contains objects that possibly belong in X. This constitutes the concept of rough classification (Pawlak, 1984). The accuracy of the approximation, $\alpha_A(X)$, is defined by Eq(1).

$$\alpha_A(X) = \frac{|\overline{Apr_A}(X)|}{\left|\underline{Apr_A}(X)\right|} \tag{1}$$

RST is also used to remove redundant information. A minimal set of condition attributes, $C' \subset C$, which maintains the classification of objects into the correct decision classes or maintains the dependency between condition and decision classes is known as the reduct. The intersection of all reducts is defined as the core.

Information obtained from the relationship between condition and decision attributes can be translated into a rule-based model consisting of IF-THEN decision rules, which can be used for knowledge-based decision support (Pawlak, 1997). IF-THEN decision rules are expressed as $\Phi \rightarrow \Psi$ and read as IF Φ THEN Ψ. Each rule approximates the underlying pattern for a cluster (or rough set) of examples in the dataset. In RSML, it is not necessary to assume a homogeneous data set. Individual rules may only apply to a subset of all examples encountered during training. Performance metrics in classical RSML thus focus on individual rules rather than the entire rule-based model; these metrics are linked to Bayesian probabilities (Pawlak, 2002). The certainty factor, coverage factor, and strength are used to characterize and analyze decision rules. The certainty factor, $cer_S(\Phi,\Psi)$, is also referred to as the confidence coefficient and provides information on how likely objects will be classified into a decision class given its performance in the condition attributes (Eq(2)). The coverage factor, $cov_S(\Phi,\Psi)$, indicates the fraction of objects in a decision class which were classified by the decision rule (Eq(3)). Strength, $\sigma_S(\Phi,\Psi)$, refers to the ratio between all objects classified by a decision rule and all objects in the information table (Eq(4)). Finally, the performance of the rules during validation is evaluated using Eq(5).

$$cer_S(\Phi, \Psi) = \frac{card(\|\Phi \wedge \Psi\|_S)}{card\|\Phi\|_S} \tag{2}$$

$$cov_S(\Phi, \Psi) = \frac{card(\|\Phi \wedge \Psi\|_S)}{card\|\Psi\|_S} \tag{3}$$

$$\sigma_S(\Phi, \Psi) = \frac{supp_S(\Phi, \Psi)}{card(U)} \tag{4}$$

$$\%Correct = \frac{number\ of\ majority\ correct\ predictions}{number\ of\ entries} \times 100 \tag{5}$$

4. Case Study

The case study intends to determine the relationship between an airline company's ESG performance indicators with its financial performance based on ROA. The Thomson Reuters ESG performance scores framework is summarized in Table 1. The individual category indicators, aggregate scores for the 3 pillars and for ESG were considered as condition attributes while ROA is the decision attribute.

Table 1: Thomson Reuters ESG performance score framework (Thomson Reuters, 2017; Refinitiv, 2022)

Pillar	Category
Environmental	Resource Use
	Innovation
	Emissions
Social	Workforce
	Human Rights
	Community
	Product Responsibility
Governance	Management
	Shareholders
	CSR Strategy

Data was collected from the Thomson Reuters EIKON database as described in the work of Kuo et al. (2021). The data consists of the performance of 30 different airline companies during the years 2012 – 2016. This resulted in 133 samples which were considered valid for use in classical RSML processing. The data was processed using the software ROSE2 (Predki et al., 1998). To interface with the software ROSE2, performance in each indicator category were sorted and divided into quartiles, with High values in the first quartile (Q1), Average in Q2 and Q3, then Poor for those in Q4. Among those valid samples, 28 (21.05 % of samples) were set aside to serve as validation samples to rate the effectiveness of the model in predicting ROA. The remaining 105 (78.95 % of samples) were used as training data. Using this information, reducts and their corresponding rule-based models were generated. Reduct generation was performed using ROSE2's Heuristic Search function (at relative quality of 95 %, two attributes considered per iteration, and minimum quality increase of 95 %), which resulted in 4 unique reducts. All reducts then underwent model generation using ROSE2's Satisfactory Description function (at maximum rule length of three, relative minimum strength of 30 %, and minimum discretion level of 60 %). The generated rule-based models from each reduct were then recorded and identified (R_i). The performance of the rule-based models were validated against the validation data. The predictive performance of the rule was determined using Eq(5). The effectiveness of the rule-based models were recorded and analyzed to identify which performed best and to determine which were most effective in predicting ROA.

5. Results and Discussion

Four (4) reducts were generated consisting of 8 attributes from the original 14, the attributes present in all reducts being ESG, Innovation, Work Force, Human Rights, Product Responsibility, and Shareholders. These attributes constitute the Core and are indicated with an asterisk in Table 2. These are the most relevant attributes, removing any one of them can influence the classification power of the attributes. The attributes selected in each reduct are shown in Table 2. All reduct sets can classify the training data to the same extent.

Rule-based models R2 and R3 were found to be similar even when the list of attributes differed. The performance of each rule is summarized in Table 3. Rule 1 indicates that if the company's performance in resource use is poor then the company's financial performance will be high. This result is consistent with the findings obtained by Tan et al. (2017) for the airline industry, citing that companies may be hesitant in investing in activities that address environmental performance because these come at a cost with uncertain financial benefits. Companies thus choose to invest in activities linked to other pillars of ESG. Rules 2 to 5 predict an average performance in the company's ROA. The result corroborates literature that organizations' commitment to fulfil their social responsibility results in moderate to good financial performance (Waworuntu et al., 2014). For Korean firms, social contribution was the only CSR factor that positively impacted a firm's financial performance (Cho et al., 2019). The good corporate citizenship performed by protecting public health and promoting good business ethics (Refinitiv, 2022) create sustainable competitive advantage and improves public reputation (Tsang et al., 2022).

Scores for environment disclosures are well-established and are being used for a longer period (Gillian et al., 2021). This supports the inclusion of the level of resource use as a predictor (see Rules 1, 3, 4, and 5). An average rating for resource use is helpful in predicting average financial performance, despite poor rating for ESG or shareholders (Rules 3 and 5). It has been argued that the social pillar of ESG reporting is difficult to measure (Ruggie and Middleton, 2019), specifically respect for human rights (De Felice, 2015), with the scores relying on transparency weights rather than quantitative industry median (Refinitiv, 2022). If the companies' efforts are known to the public and this positive perception turns into customer satisfaction (Ghanbarpour and Gustaffson, 2021), it impacts financial performance (Rules 2 and 4). Investors see positive ESG scores as a reputation insurance that predicts future performance of the company (Tsang, 2023). ESG disclosure tends to provide more benefits for the company in the long run (beyond five years); it serves as a competitive advantage that helps improve financial performance and firm value (Rojo-Suárez and Alonso-Conde, 2023).

*Table 2: Attributes in reducts (Attributes included in the core are indicated with *)*

Condition	R1	R2	R3	R4
ESG*	✓	✓	✓	✓
Environmental				
Resource Use		✓	✓	
Innovation*	✓	✓	✓	✓
Emissions	✓			
Social				
Workforce*	✓	✓	✓	✓
Human Rights*	✓	✓	✓	✓
Community				✓
Product Responsibility*	✓	✓	✓	✓
Governance	✓	✓		✓
Management			✓	
Shareholders*	✓	✓	✓	✓
CSR Strategy				

Table 3: Rules using R2 and R3 and their performance

Rules		$supp_S$	cer_S	cov_S	σ_S
1	(Resource Use = Poor) → (ROA = High)	18	0.60	0.62	0.17
2	(Human Rights = Average) → (ROA = Average)	28	0.61	0.57	0.27
3	(ESG = Poor) & (Resource Use = Average) → (ROA = Average)	16	0.76	0.33	0.15
4	(Resource Use = Average) & (Workforce = Average) → (ROA = Average)	15	0.60	0.31	0.14
5	(Resource Use = Average) & (Shareholders = Poor) → (ROA = Average)	15	0.60	0.31	0.14

Table 4: Performance of R2 and R3 in validation data

Rules	Matched data	Correct prediction	Incorrect prediction	% Correct
1	8	2	6	25.0
2	14	10	4	71.4
3	6	3	3	50.0
4	9	5	4	55.6
5	4	3	1	75.0

The rule-based models were then used to predict the performance of airline companies included in the validation dataset. The performance of R2 and R3 are shown in Table 4. Only 26 of 28 samples matched at least one rule. Over-all, R2 and R3 were able to correctly predict the ROA of 15 out of the 28-validation data (53.57 %). 12 out of the 28 (42.86 %) were incorrectly classified. Due to poor performance, Rule 1 can be removed without affecting the other rules. Of particular interest are Rules 2 and 5. Rule 2 had the highest support both in the training and validation data, with a prediction capability of 71.4 %. Rule 5 had the highest prediction capability. Examining all the rules generated by the 4 reducts reveals that there are only 8 unique rules, which are no longer presented for brevity. Using Eq(5), the over-all performance of the rule–based models in predicting ROA using the validation data is shown in Figure 1. Rule–based models R2 and R3 performed best in terms of coverage and prediction capability. R4 had the lowest percentage of incorrectly classified samples (25 %) but also had the lowest coverage, covering only 67.86 % of the validation data. More analysis is needed in examining the trade-offs between prediction performance and rule coverage.

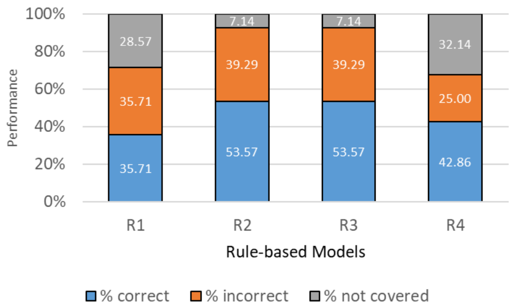

Figure 1: Prediction performance of rules from reducts

6. Conclusions

A rule-based model using RSML for predicting the financial performance of the airline industry is developed in this work using data obtained from the Thomson Reuters EIKON database. The model generated only predicted for high and average ROA and performed better in predicting the performance of companies with average ROA than those with high ROA. The best performing rule–based model had an overall predictive capability of 53.57 %. The rule with the highest coverage (> 50 % of training and validation data) which pertained to the relation between the Human Rights indicator and ROA performance had a predictive capability of 71.4 %, emphasizing the relevance of this social attribute in a company’s value. This work demonstrates how RSML can be used to generate empirical rules to predict firm ESG performance from corporate data. Future work can look at further improving the performance of the model by including additional factors such as the influence of ownership type within the model, the type of airline, as well as a longer coverage on financial performance. The approach can also be extended to other types of industries. Controversies may also be considered, as these account for organizations’ law violations and public perception issues.

Nomenclature

$\alpha_A(X)$ – accuracy of approximation
σ_S – strength of a rule
$\underline{Apr}_A(X)$ – lower approximation of X
$\overline{Apr}_A(X)$ – upper approximation of X
card – cardinality
$\mathrm{cer}_S(\Phi, \Psi)$ – certainty of a rule
$\mathrm{cov}_S(\Phi, \Psi)$ – coverage of a rule
supp_S – support of a rule

References

Amankwah-Amoah J., Debrah Y., Anang S., 2023, Greening aviation in era of COVID-19: Towards conceptualizing and operationalizing decarbonization, Journal of Environmental Management, 326, 116649.

Aviso K.B., Capili M.J., Chin H.H., Fan Y.V., Klemeš J.J., Tan R.R., 2021, Detecting patterns in energy use and greenhouse gas emissions of cities using machine learning, Chemical Engineering Transactions, 88, 403–408.

Aydoğmuş M., Gülay G., Ergun K., 2022, Impact of ESG performance on firm value and profitability, Borsa Istanbul Review, 22, S119–S127.

Bang J., Ryu D., Yu J., 2023, ESG controversies and investor trading behavior in the Korean market, Finance Research Letters, 54, 103750.

Bergero C., Gosnell G., Gielen D., Bazilian M., Davis S.J., 2023, Pathways to net-zero emissions from aviation, Nature Sustainability, 6, 404–414.

Cho S., Chung C.Y., Young J.C., 2019, Study on the relationship between CSR and financial performance, Sustainability, 11, 343.

De Felice D., 2015, Business and human rights indicators to measure the corporate responsibility to respect: challenges and opportunities, Human Rights Quarterly, 37, 511–555.

Fahey D.W., Lee D.S., 2016, Aviation and Climate Change: A Scientific Perspective. Carbon & Climate Law Review, 10(2), 97–104.

Gillan S.L., Koch A.M., Starks L.T., 2021, Firms and social responsibility: A review of ESG and CSR research in corporate finance. Journal of Corporate Finance, 66, 101889.

Ghanbarpour T., Gustafsson A., 2021, How do corporate social responsibility (CSR) and innovativeness increase financial gains? A customer perspective analysis, Journal of Business Research, 140, 11016.

Gupta A., Walton R.O., 2017, Assessment of air cargo airlines: An interpretive structural modeling approach, Procedia Manufacturing, 11, 1908−1915.

IPCC, 2022, Summary for Policymakers. In: Climate Change 2022: Mitigation of Climate Change. Contribution of Working Group III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press. Cambridge, UK and New York, NY, USA.

Khan M.A., Abbas K., Su'ud M.M., Salameh A.A., Alam M.M., Aman N., Mehreen M., Jan A., Hashim N.A.A.B.N., Aziz R.C., 2022, Application of machine learning algorithms for sustainable business management based on macro-economic data: Supervised learning techniques approach, Sustainability, 14, 9964.

Kuo T.C., Chen H.M., Meng H.M., 2021, Do corporate social responsibility practices improve financial performance? A case study of airline companies, Journal of Cleaner Production, 310, 127380.

Pawlak Z., 1982, Rough sets, International Journal of Computer & Information Sciences, 11, 341−356.

Pawlak Z., 1984, Rough classification, International Journal of Man-Machine Studies, 20, 469−483.

Pawlak Z., 1997, Rough set approach to knowledge-based decision support, European journal of operational research, 99, 48−57.

Pawlak Z., 2002, Rough sets, decision algorithms and Bayes' theorem, European Journal of Operational Research, 136, 181–189.

Perryman M., Besco L., Suleiman C., Lucato L., 2022, Ready for take off: Airline engagement with the United Nations Sustainable Development Goals, Journal of Air Transport Management, 103, 102246.

Predki B., Slowinski R., Stefanowski J., Susmaga R., Wilk S., 1998, ROSE - Software implementation of the rough set theory. In: L. Polkowski, A. Skowron (Eds.), Rough sets and current trends in computing, Lecture notes in artificial intelligence, Vol. 1424, Springer-Verlag, Berlin, Germany, 605–608.

Refinitiv, 2022. Environmental, social and governance scores from Refinitiv. <www.refinitiv.com/content/dam/marketing/en_us/documents/methodology/refinitiv-esg-scores-methodology.pdf> 09.04.2019

Rojo-Suárez J., Alonso-Conde A.B., 2023, Short-run and long-run effects of ESG policies on value creation and the cost of equity of firms, Economic Analysis and Policy, 77, 599–616.

Rudin C., 2019. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nature Machine Intelligence, 1, 206−215.

Ruggie J.G., Middleton E.R., 2019, Money, Millennials and Human Rights: Sustaining 'Sustainable Investing,' Global Policy, 10, 144–150.

Tan, S.H., Habibullah, M.S., Tan, S.K., Choon, S.W., 2017. The impact of the dimensions of environmental performance on firm performance in travel and tourism industry. Journal of environmental management, 203, 603-611.

Thomson Reuters, 2017. Thomson Reuters ESG Scores. Available online at <www.esade.edu/itemsweb/biblioteca/bbdd/inbbdd/archivos/Thomson_Reuters_ESG_Scores.pdf> accessed on 11.05.2023.

Torres L., Ripa D., Jain A., Herrero J., Leka S., 2023, The potential of responsible business to promote sustainable work – an analysis of CSR/ESG Instruments, Safety Science, 164, 106151.

Tsang A., Frost T.S., Cao H., 2022, Environmental, Social, and Governance (ESG) disclosure: A literature review, British Accounting Review, 55, 101149.

Waworuntu S.R., Wantah M.D., Rusmanto T., 2014, CSR and Financial Performance Analysis: Evidence from Top ASEAN Listed Companies. Procedia − Social and Behavioral Sciences, 164, 493–500.

Yang A.S., Baasandorj S., 2017, Exploring CSR and financial performance of full-service and low-cost air carriers, Finance Research Letters, 23, 291–299.

Zieba M., Johansson E., 2022, Sustainability reporting in the airline industry: Current literature and future research avenues, Transportation Research Part D − Transport and Environment, 102, 103133.