Supporting Information for Scope 3 emissions: data quality and machine learning prediction accuracy

S4 Table. The top five most relevant predictors in stepwise regression

Predictors	Total Scope 3	[1] Purchased Goods and Service	[2] Capital Goods	[3] Fuel and Energy Related Activities	[4] Upstream Transportation and Distribution	[5] Waste Generation	[6] Business Travel	[7] Employee Commuting	[8] Upstream Leased Assets	[9] Downstream Transportation and Distribution	[10] Process of Sold Products	[11] Use of Sold Products	[12] End of Line Treatment	[13] Downstream Leased Assets	[14] Franchise	[15] Investments	[16] Other
Revenue	1	2		4	1	4	2	1	1	4		1	1				4
Number of Employees			3				1	2									
Gross Profit							5	4									
Operational Expense								5							2		
Net Property, Plant and Equipment											1	1			1		
Capital Expense	3		1	1		1											1
Leverage														1			
Asset Age			4														
Total Debt															2		
Receivable			2													5	

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Inventory	1		3	2							2	3	3			2
Gross Intangible Asset						3	3									
Current Asset	3															
GMAR											5					
Current Liability							5									
[GICS] Banks															4	
[GICS] Commercial & Professional Services																
[GICS] Consumer Services													1			
[GICS] Diversified Financials																3
[GICS] Energy		5							3	3			4			5
[GICS] Food, Beverage & Tobacco	5			5					2							
[GICS] Household & Personal Products												4				
[GICS] Insurance								2								
[GICS] Materials	4		5	3	5			5	3	2		2		3	3	
[GICS] Pharmaceuticals, Biotechnology & Life Sciences														5		
[GICS] Retailing								3								
[GICS] Software & Services					3	4						5				
[GICS] Telecommunication Services																
[GICS] Transportation								4					5			
[GICS] Utilities			2								4					

This table summarises the top five most relevant predictors for each emissions scope in *Forward-Backward Stepwise Regression*, which automatically includes relevant predictors (< 1% significant level) into the model and excludes irrelevant ones (>5% significant level) as described in Section 4.3.2. Rank of each predictor is determined by the order of which the predictors are included into the predictor set.