Engineered Features – From Raw Data to Model Format

No.	Feature Name	Original Table(s)	Raw Data Sample (Before Engineering)	Model Type	Transformation Logic	Sample Values
1	org_young	charity	Registration Date: 2016-08-15 → Age ≈ 8 years	Binary	Age between 4–8 years → 1 else 0	0, 1
2	env_diversified	classification	Classifications: ["EDU", "POV", "HLTH", "COMM"]	Binary	If charity has ≥4 unique categories → 1	0, 1
3	fin_single_source	ar_parta or annual_return	Grants: £9,000, Donations: £1,000 (Total: £10,000)	Binary	If ≥90% of income from single source → 1	0, 1
4	temp_recent_activity	annual_return	Returns filed on: 2023-04-01, 2021-03-28	Binary	If return submitted in past 2 years → 1	0, 1
5	text_brief	charity	Activity: "Supports local causes and projects"	Binary	If activity description <100 characters → 1	0, 1
6	text_clear_mission	charity	Activity: "We serve the vulnerable communities in London"	Binary	If ≥3 clarity keywords (e.g., "support", "serve", etc.) → 1	0, 1
7	org_strong_governance	trustee	12 rows → 12 trustees	Binary	Trustee count >10 → 1	0, 1
8	org_adequate_governance	trustee	7 rows → 7 trustees	Binary	5 ≤ Trustee count ≤ 10 → 1	0, 1
9	fin_small_charity	ar_parta / annual_return	Income: £8,500	Binary	If income < £10,000 → 1	0, 1
10	org_very_young	charity	Registration Date: 2023-01-10 (Age: 2 years)	Binary	Age ≤ 3 years → 1	0, 1

11	org_broad_scope	classification	Classifications: ["HLTH", "YOUTH", "EDU", "POV", "REL"]	Binary	If org spans >3 classifications → 1	0, 1
12	fin_large_charity	ar_parta / annual_return	Income: £1,800,000	Binary	If income > £1,000,000 → 1	0, 1
13	temp_compliant	annual_return	Returned filed: 2020, 2021, 2022 — all on time	Binary	If all recent returns filed on time → 1	0, 1
14	org_mature	charity	Registration: 1990-05-01 (Age ≈ 34 years)	Binary	If age ≥ 20 years → 1	0, 1
15	fin_basic_efficiency	charity	Website: www.org.org, Email: info@org.org, Phone: 01234 567 890	Binary	If any of [email, website, phone] present → 1	0, 1