

1 Multirow Test

1.1 Schema Statistics

Schema	Domain	Total Tables	Total Columns	Total Foreign Keys	Total Primary Keys
Sakila	Entertainment	16	90	22	12
IMDB	Entertainment	7	39	7	2
MIMIC	Healthcare	26	324	55	6
CPRD Aurum	Healthcare	8	76	21	5
CPRD Gold	Healthcare	9	123	21	4
CMS	Healthcare	5	96	4	1
OMOP	Healthcare	39	432	176	12

1.2 Performance

Method	IMDB-Sakila			OMOP-CMS			CPRD Aurum-OMOP			CPRD Gold-OMOP		
	P	R	F1	P	R	F1	P	R	F1	P	R	F1
Coma	0.71	0.14	0.23	1.0	0.01	0.01	0.01	0.02	0.01	0.0	0.01	0.01
Similarity	0.18	0.22	0.2	0.38	0.26	0.3	0.03	0.15	0.05	0.01	0.03	0.01
Flooding												
Unicorn	0.986	0.0	0.0	0.967	0.002	0.003	0.980	0.045	0.018	0.995	0.089	0.035
Rematch	0.29	0.890	0.43	0.31	<u>0.770</u>	0.45	0.03	0.730	0.05	0.01	<u>0.400</u>	0.02
(gpt-3.5)												
Rematch	0.41	<u>0.830</u>	0.55	0.3	<u>0.770</u>	0.44	0.16	0.5	0.24	0.03	0.32	0.05
(gpt-4o)												
Schema Un- derstanding	0.77	0.67	<u>0.720</u>	0.41	0.800	<u>0.540</u>	<u>0.450</u>	0.5	0.480	0.17	0.32	<u>0.220</u>
(gpt-3.5)												
Schema Un- derstanding	<u>0.900</u>	0.75	0.820	<u>0.750</u>	0.51	0.610	0.3	<u>0.670</u>	<u>0.420</u>	<u>0.190</u>	0.410	0.260
(gpt-4o)												