Failures caused mainly by errors in textual attributes

Legend: GPT-4o-Mini, GPT-4o, Claude 3.5 Sonnet, Llama 3.1 Instruct

	Clean	Deduction ≤ $0.1	Assignment Number	Billing Number	Partner Name
GPT-4o-Mini	0.96	0.93	0.89	0.91	0.14
GPT-4o	0.97	0.97	0.93	0.91	0.57
Claude 3.5 Sonnet	1.00	0.99	0.98	0.97	0.80
Llama 3.1 Instruct	0.96	0.96	0.91	0.88	0.83

Y-axis: F1 Score