Overall Performance Heatmap: Models vs Experiments

Model	Exp 1	Exp 2	Exp 3	Exp 4	Exp 5	Exp 6	Exp 7
GPT OSS 120B	0.774	0.701	0.766	0.842	0.847	0.757	0.816
GPT OSS 20B	0.815	0.701	0.870	0.801	0.791	0.779	0.785
GPT 4 Nano	0.781	0.861	0.745	0.791	0.791	0.664	0.733