

Model	Size	HumanEval		MBPP	DS-1000
		Python	Multilingual		
Pre-Trained Model					
Codex-001	-	33.5	26.1	45.9	20.2
Codex-002	-	-	-	-	39.2
CodeGeeX2	6B	36.0	24.5	42.4	22.9
StarCoder	16B	36.0	28.7	46.8	27.2
CodeLlama	7B	31.7	29.2	41.6	22.1
	13B	36.0	35.4	48.4	26.8
	34B	48.2	41.0	55.2	34.3
DeepSeek-Coder Base	1B	34.8	28.3	46.2	16.2
	7B	49.4	44.7	60.6	30.5
	33B	56.1	50.3	66.0	40.2
Instruction-Tuned Model					
GPT-3.5-turbo	-	76.2	64.9	70.8	-
GPT4	-	84.1	76.5	80.0	-
DeepSeek-Coder Instruct	7B	78.6	66.1	65.4	-
	33B	79.3	69.2	70.0	-