0.7

0.6

0.5

0.4

0.3

0.2

0.1

GPT2-124M (backwards)-	0.68						
GPT2-355M-	0.77	0.64					
GPT2-355M (backwards)-	0.64	0.68	0.71				
GPT2-774M -	0.62	0.58	0.84	0.73			
GPT2-774M (backwards)-	0.59	0.66	0.69	0.73	0.73		
Human experts-	0.11	0.029	0.15	0.076	0.12	0.045	
	GPT2-124M-	GPT2-124M (backwards)	GPT2-355M-	GPT2-355M (backwards)	GPT2-774M-	GPT2-774M (backwards)	