
Supplementary Material of “*Magical*: Medical Lay Language Generation via Semantic Invariance and Layperson-tailored Adaptation”

Anonymous Author(s)

Affiliation

Address

email

1 In this supplementary material, we conduct statistical experiments on the results reported in our
2 main table to evaluate the significance of the performance improvements of *Magical* over recent
3 state-of-the-art methods. Specifically, we employ the bootstrap method, randomly sampling 95% of
4 the test set for evaluation and repeating this process 100 times. The mean and standard deviation
5 across these repeated experiments are reported in Table. 1. The results demonstrate that *Magical*
6 achieves statistically significant improvements compared to recent works in the literature.

Table 1: Performance comparison of Prompt, various LoRA variants, and *Magical* across three backbone LLMs on three MLLG datasets. All evaluation metrics are defined such that higher values indicate better performance (†). In each group of experiments, the value before the parentheses denotes the mean, while the value inside the parentheses denotes the standard deviation.

Methods	#Params	Cochrane				eLife				Plos_genetics			
		R-1	R-2	R-L	BLEU	R-1	R-2	R-L	BLEU	R-1	R-2	R-L	BLEU
LLaMA3.2-3B-Instruct													
Prompt	N/A	39.80 (0.054)	11.37 (0.042)	36.51 (0.051)	5.29 (0.037)	37.82 (0.082)	8.22 (0.047)	35.47 (0.083)	3.07 (0.033)	37.89 (0.079)	7.78 (0.049)	35.08 (0.074)	4.13 (0.037)
LoRA	36M	39.98 (0.155)	16.31 (0.122)	37.68 (0.151)	10.35 (0.127)	46.69 (0.127)	13.47 (0.071)	44.57 (0.121)	7.35 (0.072)	44.91 (0.087)	13.63 (0.086)	41.42 (0.081)	8.80 (0.096)
rsLoRA	36M	40.87 (0.193)	16.38 (0.131)	38.46 (0.184)	10.23 (0.177)	45.41 (0.132)	13.02 (0.080)	43.30 (0.131)	7.10 (0.059)	43.77 (0.137)	13.92 (0.090)	40.55 (0.131)	7.81 (0.099)
DoRA	37M	40.07 (0.184)	16.29 (0.128)	37.85 (0.178)	10.29 (0.161)	42.98 (0.167)	12.03 (0.069)	40.89 (0.159)	6.17 (0.047)	41.15 (0.148)	12.80 (0.099)	38.42 (0.138)	6.44 (0.088)
PISSA	36M	39.72 (0.182)	16.14 (0.127)	37.36 (0.172)	10.51 (0.166)	42.88 (0.122)	12.00 (0.069)	40.85 (0.118)	6.26 (0.049)	43.53 (0.141)	13.93 (0.089)	40.47 (0.136)	7.80 (0.099)
Magical	24M	45.31 (0.164)	19.39 (0.150)	42.34 (0.164)	16.65 (0.190)	49.14 (0.061)	14.67 (0.050)	46.89 (0.063)	8.29 (0.049)	47.50 (0.105)	15.46 (0.087)	44.02 (0.102)	10.22 (0.095)
LLaMA3.1-8B-Instruct													
Prompt	N/A	41.68 (0.057)	12.50 (0.047)	38.61 (0.058)	6.22 (0.044)	35.83 (0.082)	9.45 (0.043)	33.37 (0.076)	1.86 (0.024)	39.13 (0.078)	8.66 (0.043)	36.16 (0.076)	4.66 (0.042)
LoRA	62M	40.25 (0.170)	16.59 (0.114)	37.57 (0.168)	12.11 (0.174)	49.40 (0.063)	14.88 (0.048)	47.29 (0.065)	8.53 (0.041)	47.56 (0.123)	16.34 (0.102)	43.98 (0.115)	11.53 (0.124)
rsLoRA	62M	43.30 (0.132)	18.11 (0.117)	40.62 (0.128)	12.89 (0.160)	49.33 (0.097)	15.00 (0.075)	47.26 (0.094)	8.70 (0.057)	42.18 (0.127)	12.86 (0.082)	39.36 (0.117)	6.87 (0.071)
DoRA	64M	43.24 (0.143)	18.29 (0.116)	40.60 (0.142)	13.49 (0.162)	48.47 (0.095)	14.66 (0.059)	46.40 (0.093)	8.43 (0.044)	42.47 (0.139)	13.04 (0.089)	39.58 (0.130)	7.17 (0.078)
PISSA	62M	42.94 (0.146)	17.80 (0.118)	40.22 (0.141)	13.91 (0.180)	48.81 (0.102)	14.64 (0.074)	46.73 (0.101)	8.38 (0.051)	39.63 (0.155)	11.62 (0.097)	37.25 (0.146)	5.62 (0.078)
Magical	42M	45.74 (0.153)	19.54 (0.144)	42.82 (0.158)	16.70 (0.191)	50.43 (0.071)	15.48 (0.053)	48.01 (0.068)	8.66 (0.043)	48.77 (0.084)	16.64 (0.089)	45.06 (0.080)	11.85 (0.088)
Qwen2.5-7B-Instruct													
Prompt	N/A	44.52 (0.097)	15.56 (0.080)	41.51 (0.094)	11.41 (0.075)	35.10 (0.098)	8.86 (0.049)	32.29 (0.093)	1.80 (0.024)	44.93 (0.092)	12.50 (0.060)	41.13 (0.092)	7.72 (0.053)
LoRA	60M	44.20 (0.131)	19.35 (0.114)	41.57 (0.134)	14.48 (0.148)	47.37 (0.107)	13.89 (0.057)	45.41 (0.104)	7.93 (0.045)	46.39 (0.138)	16.29 (0.093)	42.92 (0.131)	9.93 (0.113)
rsLoRA	60M	45.23 (0.125)	19.48 (0.107)	42.46 (0.125)	15.06 (0.161)	48.72 (0.102)	14.54 (0.064)	46.50 (0.102)	7.76 (0.054)	46.26 (0.155)	15.62 (0.105)	42.87 (0.148)	9.92 (0.110)
DoRA	61M	45.68 (0.135)	19.66 (0.108)	42.90 (0.132)	15.47 (0.152)	47.36 (0.104)	13.90 (0.066)	45.22 (0.102)	7.70 (0.053)	46.67 (0.149)	15.99 (0.098)	43.27 (0.142)	10.13 (0.114)
PISSA	60M	44.81 (0.156)	19.26 (0.140)	42.07 (0.156)	14.46 (0.193)	48.86 (0.099)	14.46 (0.060)	46.68 (0.100)	7.83 (0.043)	46.81 (0.117)	15.92 (0.092)	43.57 (0.108)	10.48 (0.088)
Magical	42M	47.53 (0.132)	20.85 (0.133)	44.52 (0.133)	17.33 (0.156)	50.47 (0.067)	14.97 (0.053)	48.17 (0.066)	8.41 (0.041)	48.54 (0.101)	16.41 (0.090)	44.78 (0.095)	11.50 (0.082)