

Table 17 Vocabulary Diversity by Model, Sampling Strategy, and Task Category.

(Continued from Table 16)

Model	Sampling Strategy	A	B	C	D	E	F	G	H
Llama-3.1-8B-Instruct	Temperature (t=0.0)	0.15 (0.03)	0.07 (0.02)	0.24 (0.06)	0.45 (0.03)	0.54 (0.01)	0.47 (0.04)	0.46 (0.04)	0.47 (0.02)
Llama-3.1-8B-Instruct	Temperature (t=0.5)	0.43 (0.03)	0.32 (0.05)	0.48 (0.09)	0.64 (0.02)	0.66 (0.01)	0.63 (0.02)	0.65 (0.03)	0.67 (0.01)
Llama-3.1-8B-Instruct	Temperature (t=1.0)	0.64 (0.02)	0.57 (0.04)	0.74 (0.05)	0.72 (0.01)	0.74 (0.00)	0.74 (0.02)	0.77 (0.02)	0.75 (0.01)
Llama-3.1-8B-Instruct	In-Context Regeneration (General)	0.58 (0.02)	0.86 (0.05)	0.94 (0.03)	0.53 (0.02)	0.52 (0.02)	0.64 (0.04)	0.73 (0.03)	0.68 (0.02)
Llama-3.1-8B-Instruct	In-Context Regeneration (Task-Anchored)	0.52 (0.02)	0.89 (0.04)	0.87 (0.07)	0.60 (0.01)	0.66 (0.01)	0.73 (0.03)	0.84 (0.02)	0.80 (0.01)
Llama-3.1-8B-Instruct	System Prompt (General)	0.65 (0.02)	0.73 (0.03)	0.72 (0.04)	0.42 (0.02)	0.64 (0.02)	0.66 (0.04)	0.71 (0.03)	0.76 (0.01)
Llama-3.1-8B-Instruct	System Prompt (Task-Anchored)	0.58 (0.02)	0.66 (0.03)	0.49 (0.06)	0.57 (0.03)	0.68 (0.01)	0.72 (0.03)	0.76 (0.02)	0.76 (0.01)
Mistral-7B-Instruct-v0.3	Temperature (t=0.0)	0.33 (0.03)	0.26 (0.06)	0.38 (0.06)	0.49 (0.03)	0.55 (0.01)	0.50 (0.04)	0.53 (0.04)	0.46 (0.02)
Mistral-7B-Instruct-v0.3	Temperature (t=0.5)	0.60 (0.02)	0.60 (0.03)	0.74 (0.02)	0.70 (0.01)	0.67 (0.01)	0.67 (0.02)	0.72 (0.02)	0.67 (0.01)
Mistral-7B-Instruct-v0.3	Temperature (t=1.0)	0.72 (0.01)	0.75 (0.01)	0.82 (0.02)	0.76 (0.01)	0.73 (0.01)	0.75 (0.01)	0.80 (0.01)	0.75 (0.01)
Mistral-7B-Instruct-v0.3	In-Context Regeneration (General)	0.46 (0.02)	0.73 (0.02)	0.63 (0.06)	0.27 (0.03)	0.28 (0.02)	0.43 (0.04)	0.51 (0.03)	0.53 (0.02)
Mistral-7B-Instruct-v0.3	In-Context Regeneration (Task-Anchored)	0.42 (0.02)	0.59 (0.04)	0.47 (0.08)	0.31 (0.03)	0.38 (0.02)	0.51 (0.04)	0.63 (0.03)	0.59 (0.02)
Mistral-7B-Instruct-v0.3	System Prompt (General)	0.64 (0.02)	0.80 (0.02)	0.81 (0.03)	0.45 (0.03)	0.72 (0.02)	0.74 (0.02)	0.72 (0.03)	0.80 (0.01)
Mistral-7B-Instruct-v0.3	System Prompt (Task-Anchored)	0.63 (0.02)	0.77 (0.02)	0.79 (0.02)	0.52 (0.03)	0.76 (0.02)	0.81 (0.01)	0.78 (0.02)	0.84 (0.01)

Table 18 Embedding Diversity by Model, Sampling Strategy, and Task Category.

Model	Sampling Strategy	A	B	C	D	E	F	G	H
gpt-4o	Temperature (t=0.0)	0.02 (0.00)	0.03 (0.01)	0.02 (0.01)	0.01 (0.00)	0.01 (0.00)	0.02 (0.00)	0.03 (0.00)	0.02 (0.00)
gpt-4o	Temperature (t=1.0)	0.04 (0.00)	0.04 (0.01)	0.07 (0.01)	0.03 (0.00)	0.03 (0.00)	0.04 (0.00)	0.06 (0.01)	0.04 (0.00)
gpt-4o	Temperature (t=2.0)	0.05 (0.00)	0.06 (0.01)	0.07 (0.02)	0.03 (0.00)	0.03 (0.00)	0.04 (0.00)	0.07 (0.01)	0.04 (0.00)
gpt-4o	In-Context Regeneration (General)	0.07 (0.01)	0.15 (0.01)	0.16 (0.01)	0.04 (0.00)	0.04 (0.00)	0.05 (0.01)	0.10 (0.01)	0.09 (0.01)
gpt-4o	In-Context Regeneration (Task-Anchored)	0.03 (0.00)	0.15 (0.01)	0.16 (0.01)	0.03 (0.00)	0.05 (0.00)	0.06 (0.01)	0.13 (0.01)	0.11 (0.01)
gpt-4o	System Prompt (General)	0.05 (0.00)	0.14 (0.01)	0.13 (0.01)	0.03 (0.00)	0.06 (0.00)	0.08 (0.01)	0.12 (0.01)	0.11 (0.00)
gpt-4o	System Prompt (Task-Anchored)	0.02 (0.00)	0.13 (0.01)	0.12 (0.01)	0.06 (0.00)	0.08 (0.00)	0.10 (0.01)	0.14 (0.01)	0.12 (0.00)
claude-4-sonnet	Temperature (t=0.0)	0.01 (0.00)	0.01 (0.00)	0.01 (0.00)	0.01 (0.00)	0.01 (0.00)	0.01 (0.00)	0.02 (0.00)	0.01 (0.00)
claude-4-sonnet	Temperature (t=0.5)	0.02 (0.00)	0.01 (0.00)	0.02 (0.00)	0.01 (0.00)	0.02 (0.00)	0.02 (0.00)	0.04 (0.00)	0.02 (0.00)
claude-4-sonnet	Temperature (t=1.0)	0.03 (0.00)	0.02 (0.00)	0.05 (0.01)	0.01 (0.00)	0.02 (0.00)	0.02 (0.00)	0.05 (0.00)	0.03 (0.00)
claude-4-sonnet	In-Context Regeneration (General)	0.07 (0.00)	0.13 (0.01)	0.14 (0.01)	0.04 (0.00)	0.06 (0.00)	0.07 (0.01)	0.11 (0.01)	0.10 (0.00)
claude-4-sonnet	In-Context Regeneration (Task-Anchored)	0.03 (0.00)	0.12 (0.01)	0.14 (0.01)	0.06 (0.00)	0.08 (0.00)	0.09 (0.01)	0.16 (0.01)	0.13 (0.00)
claude-4-sonnet	System Prompt (General)	0.05 (0.00)	0.16 (0.01)	0.12 (0.01)	0.04 (0.00)	0.07 (0.00)	0.07 (0.01)	0.12 (0.01)	0.10 (0.00)
claude-4-sonnet	System Prompt (Task-Anchored)	0.04 (0.00)	0.16 (0.01)	0.11 (0.01)	0.06 (0.00)	0.08 (0.00)	0.08 (0.01)	0.14 (0.01)	0.12 (0.00)
gemini-2.5-flash	Temperature (t=0.0)	0.01 (0.00)	0.01 (0.01)	0.02 (0.01)	0.01 (0.00)	0.01 (0.00)	0.02 (0.00)	0.02 (0.00)	0.01 (0.00)
gemini-2.5-flash	Temperature (t=1.0)	0.04 (0.01)	0.04 (0.01)	0.09 (0.02)	0.02 (0.00)	0.03 (0.00)	0.04 (0.00)	0.07 (0.01)	0.04 (0.00)
gemini-2.5-flash	Temperature (t=2.0)	0.05 (0.01)	0.05 (0.01)	0.10 (0.01)	0.02 (0.00)	0.04 (0.00)	0.05 (0.01)	0.07 (0.01)	0.05 (0.00)
gemini-2.5-flash	In-Context Regeneration (General)	0.08 (0.00)	0.14 (0.01)	0.15 (0.01)	0.03 (0.00)	0.04 (0.00)	0.05 (0.01)	0.10 (0.01)	0.09 (0.01)
gemini-2.5-flash	In-Context Regeneration (Task-Anchored)	0.04 (0.00)	0.14 (0.01)	0.15 (0.01)	0.04 (0.00)	0.04 (0.00)	0.06 (0.01)	0.13 (0.01)	0.10 (0.00)
gemini-2.5-flash	System Prompt (General)	0.04 (0.00)	0.13 (0.01)	0.13 (0.01)	0.03 (0.00)	0.05 (0.00)	0.07 (0.01)	0.11 (0.01)	0.10 (0.00)
gemini-2.5-flash	System Prompt (Task-Anchored)	0.02 (0.00)	0.13 (0.01)	0.12 (0.01)	0.05 (0.00)	0.07 (0.00)	0.10 (0.01)	0.14 (0.01)	0.11 (0.00)

Table 19 Embedding Diversity by Model, Sampling Strategy, and Task Category.

(Continued from Table 18)

Model	Sampling Strategy	A	B	C	D	E	F	G	H
Llama-3.1-8B-Instruct	Temperature (t=0.0)	0.01 (0.00)	0.01 (0.00)	0.03 (0.01)	0.02 (0.00)	0.02 (0.00)	0.02 (0.00)	0.03 (0.00)	0.02 (0.00)
Llama-3.1-8B-Instruct	Temperature (t=0.5)	0.03 (0.00)	0.04 (0.01)	0.05 (0.01)	0.04 (0.00)	0.03 (0.00)	0.03 (0.00)	0.06 (0.01)	0.04 (0.00)
Llama-3.1-8B-Instruct	Temperature (t=1.0)	0.06 (0.00)	0.08 (0.01)	0.08 (0.01)	0.04 (0.00)	0.04 (0.00)	0.05 (0.01)	0.08 (0.01)	0.06 (0.00)
Llama-3.1-8B-Instruct	In-Context Regeneration (General)	0.07 (0.00)	0.14 (0.01)	0.16 (0.01)	0.04 (0.00)	0.04 (0.00)	0.07 (0.01)	0.11 (0.01)	0.09 (0.01)
Llama-3.1-8B-Instruct	In-Context Regeneration (Task-Anchored)	0.05 (0.01)	0.12 (0.01)	0.15 (0.01)	0.05 (0.00)	0.06 (0.00)	0.09 (0.01)	0.14 (0.01)	0.11 (0.00)
Llama-3.1-8B-Instruct	System Prompt (General)	0.07 (0.00)	0.14 (0.01)	0.11 (0.01)	0.03 (0.00)	0.07 (0.00)	0.08 (0.01)	0.10 (0.01)	0.10 (0.00)
Llama-3.1-8B-Instruct	System Prompt (Task-Anchored)	0.04 (0.00)	0.12 (0.01)	0.10 (0.01)	0.06 (0.00)	0.08 (0.00)	0.11 (0.01)	0.12 (0.01)	0.11 (0.00)
Mistral-7B-Instruct-v0.3	Temperature (t=0.0)	0.03 (0.00)	0.03 (0.01)	0.03 (0.01)	0.03 (0.00)	0.02 (0.00)	0.02 (0.00)	0.03 (0.00)	0.02 (0.00)
Mistral-7B-Instruct-v0.3	Temperature (t=0.5)	0.06 (0.00)	0.07 (0.01)	0.08 (0.01)	0.05 (0.00)	0.04 (0.00)	0.04 (0.00)	0.06 (0.01)	0.04 (0.00)
Mistral-7B-Instruct-v0.3	Temperature (t=1.0)	0.08 (0.00)	0.10 (0.01)	0.10 (0.01)	0.06 (0.00)	0.04 (0.00)	0.05 (0.00)	0.08 (0.01)	0.05 (0.00)
Mistral-7B-Instruct-v0.3	In-Context Regeneration (General)	0.04 (0.00)	0.14 (0.01)	0.11 (0.01)	0.02 (0.00)	0.02 (0.00)	0.04 (0.01)	0.06 (0.01)	0.06 (0.00)
Mistral-7B-Instruct-v0.3	In-Context Regeneration (Task-Anchored)	0.03 (0.00)	0.11 (0.01)	0.08 (0.01)	0.02 (0.00)	0.03 (0.00)	0.06 (0.01)	0.09 (0.01)	0.07 (0.00)
Mistral-7B-Instruct-v0.3	System Prompt (General)	0.07 (0.00)	0.16 (0.01)	0.13 (0.01)	0.04 (0.00)	0.08 (0.00)	0.09 (0.01)	0.11 (0.01)	0.11 (0.00)
Mistral-7B-Instruct-v0.3	System Prompt (Task-Anchored)	0.06 (0.01)	0.16 (0.01)	0.14 (0.01)	0.04 (0.00)	0.10 (0.00)	0.12 (0.01)	0.12 (0.01)	0.14 (0.00)