

Table 34: **Benchmark scores for language: German (de)**. The highest and second-best scores are shown in **bold** and underlined, respectively.

	Model	INCLUDE	MMMLU	MT-AIME24	PolyMath	Average
Thinking Mode	Gemini2.5-Pro	50.0	<u>85.6</u>	<b>86.7</b>	<u>53.8</u>	69.0
	QwQ-32B	57.1	<u>83.8</u>	76.7	<u>51.0</u>	67.2
	<b>Qwen3-235B-A22B</b>	<b>71.4</b>	<b>86.0</b>	<u>83.3</u>	<b>55.4</b>	<b>74.0</b>
	<u>Qwen3-32B</u>	<u>64.3</u>	81.9	<b>86.7</b>	48.1	<u>70.2</u>
	<b>Qwen3-30B-A3B</b>	<u>64.3</u>	81.9	80.0	46.6	68.2
	<u>Qwen3-14B</u>	<u>57.1</u>	80.9	70.0	48.1	64.0
	<u>Qwen3-8B</u>	<u>64.3</u>	78.1	66.7	43.6	63.2
	<u>Qwen3-4B</u>	<u>57.1</u>	74.0	73.3	43.1	61.9
	<u>Qwen3-1.7B</u>	<u>64.3</u>	63.4	36.7	26.8	47.8
	<u>Qwen3-0.6B</u>	<u>57.1</u>	47.6	10.0	13.7	32.1
Non-thinking Mode	GPT-4o-2024-1120	57.1	<b>80.4</b>	10.0	13.5	40.2
	Gemma-3-27b-IT	57.1	<u>76.1</u>	26.7	20.2	45.0
	Qwen2.5-72B-Instruct	<u>64.3</u>	79.9	16.7	19.3	45.0
	<b>Qwen3-235B-A22B</b>	<b>71.4</b>	<b>81.7</b>	<b>40.0</b>	<b>25.9</b>	<b>54.8</b>
	<u>Qwen3-32B</u>	<u>57.1</u>	77.2	<u>30.0</u>	21.9	46.6
	<b>Qwen3-30B-A3B</b>	<u>57.1</u>	77.7	23.3	<u>25.2</u>	45.8
	<u>Qwen3-14B</u>	<u>57.1</u>	76.0	<u>30.0</u>	<u>24.5</u>	<u>46.9</u>
	<u>Qwen3-8B</u>	<u>64.3</u>	70.8	20.0	19.9	43.8
	<u>Qwen3-4B</u>	<u>64.3</u>	66.0	26.7	16.4	43.4
	<u>Qwen3-1.7B</u>	42.9	53.2	10.0	10.6	29.2
	<u>Qwen3-0.6B</u>	42.9	37.8	3.3	5.7	22.4

Table 35: **Benchmark scores for language: Thai (th)**. The highest and second-best scores are shown in **bold** and underlined, respectively.

	Model	MLogiQA	MT-AIME24	PolyMath	Average
Thinking Mode	Gemini2.5-Pro	<u>73.8</u>	<u>80.0</u>	<u>50.7</u>	<u>68.2</u>
	QwQ-32B	<b>75.0</b>	60.0	41.3	58.8
	<b>Qwen3-235B-A22B</b>	73.8	<b>86.7</b>	<b>53.6</b>	<b>71.4</b>
	<u>Qwen3-32B</u>	<u>73.8</u>	76.7	46.9	65.8
	<b>Qwen3-30B-A3B</b>	63.7	<u>80.0</u>	45.2	63.0
	<u>Qwen3-14B</u>	65.0	<u>76.7</u>	44.4	62.0
	<u>Qwen3-8B</u>	68.8	70.0	41.3	60.0
	<u>Qwen3-4B</u>	60.0	60.0	39.4	53.1
	<u>Qwen3-1.7B</u>	48.8	33.3	23.7	35.3
	<u>Qwen3-0.6B</u>	33.8	13.3	11.4	19.5
Non-thinking Mode	GPT-4o-2024-1120	52.5	10.0	11.9	24.8
	Gemma-3-27b-IT	50.0	16.7	19.0	28.6
	Qwen2.5-72B-Instruct	<u>58.8</u>	6.7	17.4	27.6
	<b>Qwen3-235B-A22B</b>	<b>61.3</b>	<u>23.3</u>	<b>27.6</b>	<b>37.4</b>
	<u>Qwen3-32B</u>	<b>61.3</b>	13.3	22.2	32.3
	<b>Qwen3-30B-A3B</b>	50.0	<b>30.0</b>	<u>22.3</u>	<u>34.1</u>
	<u>Qwen3-14B</u>	47.5	<u>23.3</u>	22.1	31.0
	<u>Qwen3-8B</u>	42.5	10.0	17.2	23.2
	<u>Qwen3-4B</u>	43.8	13.3	16.1	24.4
	<u>Qwen3-1.7B</u>	42.5	6.7	9.5	19.6
	<u>Qwen3-0.6B</u>	37.5	0.0	3.6	13.7