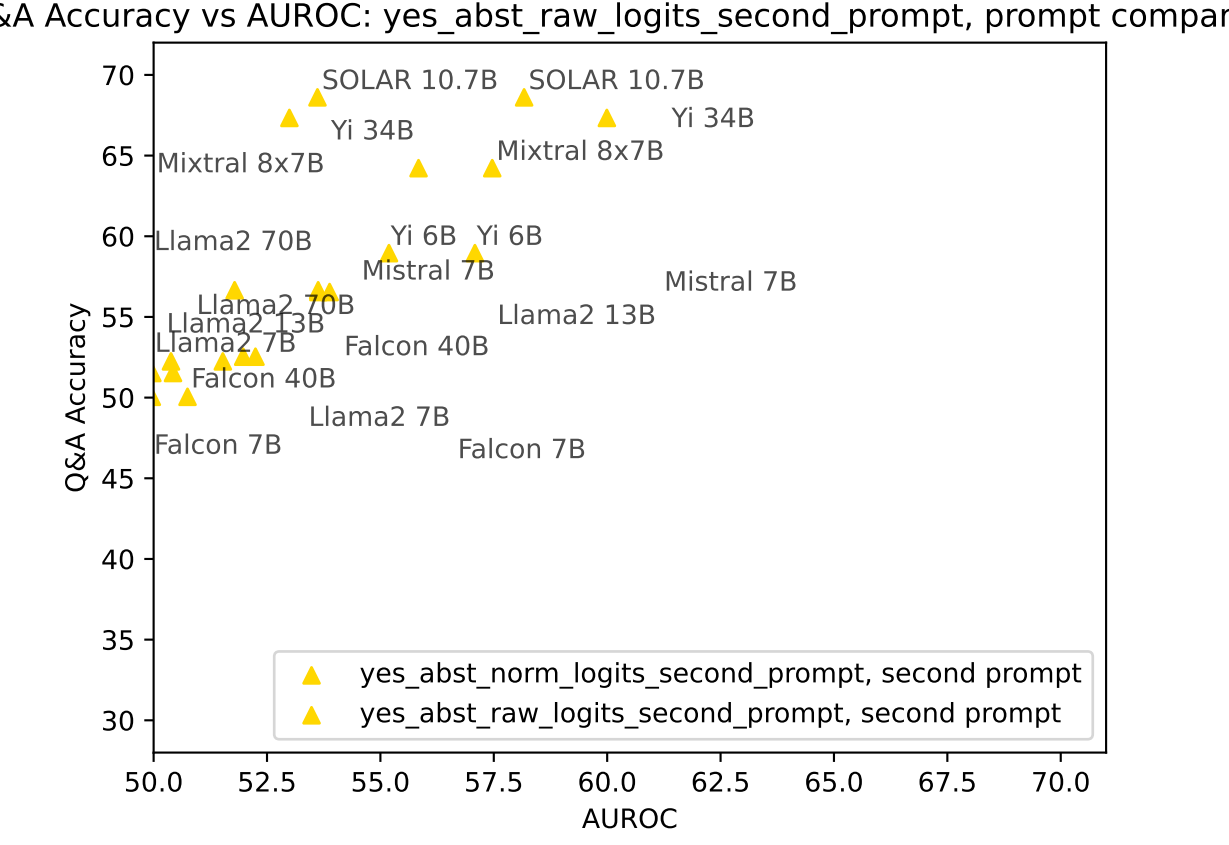
&A Accuracy vs AUROC: yes_abst_raw_logits_second_prompt, prompt compar
SOLAR 10.7B
SOLAR 10.7B
Yi 34B
Yi 34B
Mixtral 8x7B
Mixtral 8x7B
Llama2 70B
Yi 6B
Yi 6B
Mistral 7B
Mistral 7B
Llama2 70B
Llama2 13B
Llama2 13B
Llama2 7B
Falcon 40B
Falcon 40B
Llama2 7B
Falcon 7B
Falcon 7B
Q&A Accuracy
AUROC
70
65
60
55
50
45
40
35
30
50.0
52.5
55.0
57.5
60.0
62.5
65.0
67.5
70.0
yes_abst_norm_logits_second_prompt, second prompt
yes_abst_raw_logits_second_prompt, second prompt