Separability by individudal judges

Model	Separability
council (no aggregation)	92.11
council (by mean pooling)	80.53
council (by majority vote)	77.37
claude-3-opus	76.84
mistral-large	75.79
qwen1.5-110B-Chat	72.11
llama-3-70b-chat	71.58
gemini-1.5-pro	70.53
command-r-plus	69.47
gpt-4-0613	67.89
gpt-4o-2024-05-13	67.37
mistral-medium	67.37
claude-3-haiku	66.32
gpt-3.5-turbo-0125	64.21
gpt-4-turbo-2024-04-09	64.21
mixtral-8x7b	63.16
dbrx-instruct	63.16
claude-3-sonnet	63.16
command-r	58.42
gemini-1.0-pro	54.74
qwen1.5-72B-Chat	54.74
llama-3-8b-chat	45.26
qwen1.5-32B-Chat	45.26

Family:
- Council
- Anthropic
- Mistral
- Alibaba
- Meta
- Google
- Cohere
- OpenAI
- Databricks