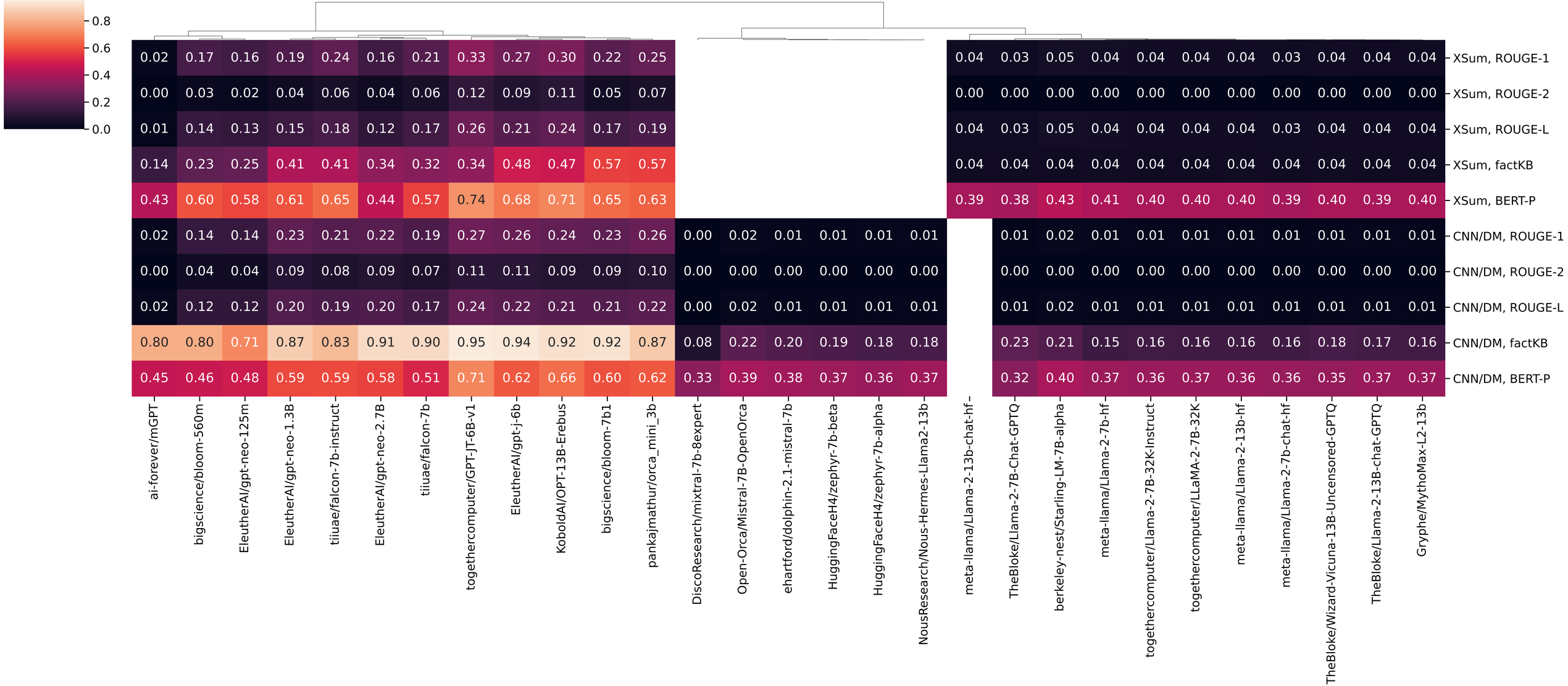
0.8
0.6
0.4
0.2
0.0
XSum, ROUGE-1 0.02 0.17 0.16 0.19 0.24 0.16 0.21 0.33 0.27 0.30 0.22 0.25 0.04 0.03 0.05 0.04 0.04 0.04 0.04 0.03 0.04 0.04 0.04
XSum, ROUGE-2 0.00 0.03 0.02 0.04 0.06 0.04 0.06 0.12 0.09 0.11 0.05 0.07 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00
XSum, ROUGE-L 0.01 0.14 0.13 0.15 0.18 0.12 0.17 0.26 0.21 0.24 0.17 0.19 0.04 0.03 0.05 0.04 0.04 0.04 0.04 0.03 0.04 0.04 0.04
XSum, factKB 0.14 0.23 0.25 0.41 0.41 0.34 0.32 0.34 0.48 0.47 0.57 0.57 0.04 0.04 0.04 0.04 0.04 0.04 0.04 0.04 0.04 0.04 0.04
XSum, BERT-P 0.43 0.60 0.58 0.61 0.65 0.44 0.57 0.74 0.68 0.71 0.65 0.63 0.39 0.38 0.43 0.41 0.40 0.40 0.40 0.39 0.40 0.39 0.40
CNN/DM, ROUGE-1 0.02 0.14 0.14 0.23 0.21 0.22 0.19 0.27 0.26 0.24 0.23 0.26 0.00 0.02 0.01 0.01 0.01 0.01 0.01 0.02 0.01 0.01 0.01 0.01 0.01 0.01 0.01 0.01 0.01
CNN/DM, ROUGE-2 0.00 0.04 0.04 0.09 0.08 0.09 0.07 0.11 0.11 0.09 0.09 0.10 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00
CNN/DM, ROUGE-L 0.02 0.12 0.12 0.20 0.19 0.20 0.17 0.24 0.22 0.21 0.21 0.22 0.00 0.02 0.01 0.01 0.01 0.01 0.01 0.02 0.01 0.01 0.01 0.01 0.01 0.01 0.01 0.01 0.01
CNN/DM, factKB 0.80 0.80 0.71 0.87 0.83 0.91 0.90 0.95 0.94 0.92 0.92 0.87 0.08 0.22 0.20 0.19 0.18 0.18 0.23 0.21 0.15 0.16 0.16 0.16 0.16 0.18 0.17 0.16
CNN/DM, BERT-P 0.45 0.46 0.48 0.59 0.59 0.58 0.51 0.71 0.62 0.66 0.60 0.62 0.33 0.39 0.38 0.37 0.36 0.37 0.32 0.40 0.37 0.36 0.37 0.36 0.36 0.35 0.37 0.37
ai-forever/mGPT
bigscience/bloom-560m
EleutherAI/gpt-neo-125m
EleutherAI/gpt-neo-1.3B
tiiuae/falcon-7b-instruct
EleutherAI/gpt-neo-2.7B
tiiuae/falcon-7b
togethercomputer/GPT-JT-6B-v1
EleutherAI/gpt-j-6b
KoboldAI/OPT-13B-Erebus
bigscience/bloom-7b1
pankajmathur/orca_mini_3b
DiscoResearch/mixtral-7b-8expert
Open-Orca/Mistral-7B-OpenOrca
ehartford/dolphin-2.1-mistral-7b
HuggingFaceH4/zephyr-7b-beta
HuggingFaceH4/zephyr-7b-alpha
NousResearch/Nous-Hermes-Llama2-13b
meta-llama/Llama-2-13b-chat-hf
TheBloke/Llama-2-7B-Chat-GPTQ
berkeley-nest/Starling-LM-7B-alpha
meta-llama/Llama-2-7b-hf
togethercomputer/Llama-2-7B-32K-Instruct
togethercomputer/LLaMA-2-7B-32K
meta-llama/Llama-2-13b-hf
meta-llama/Llama-2-7b-chat-hf
TheBloke/Wizard-Vicuna-13B-Uncensored-GPTQ
TheBloke/Llama-2-13B-chat-GPTQ
Gryphe/MythoMax-L2-13b