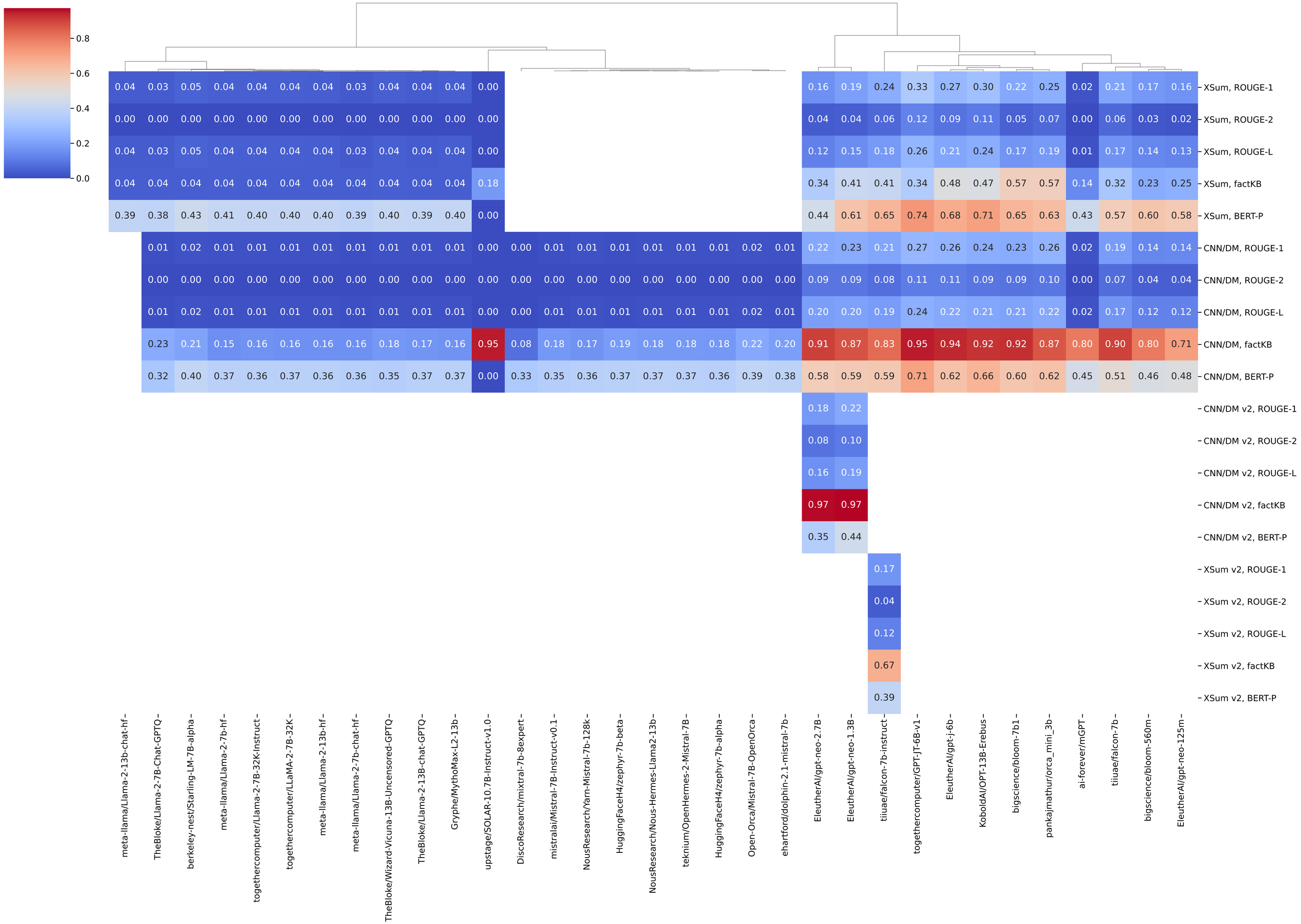
0.8
0.6
0.4
0.2
0.0
XSum, ROUGE-1 0.04 0.03 0.05 0.04 0.04 0.04 0.04 0.03 0.04 0.04 0.04 0.00 0.16 0.19 0.24 0.33 0.27 0.30 0.22 0.25 0.02 0.21 0.17 0.16
XSum, ROUGE-2 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.04 0.04 0.06 0.12 0.09 0.11 0.05 0.07 0.00 0.06 0.03 0.02
XSum, ROUGE-L 0.04 0.03 0.05 0.04 0.04 0.04 0.04 0.03 0.04 0.04 0.04 0.00 0.12 0.15 0.18 0.26 0.21 0.24 0.17 0.19 0.01 0.17 0.14 0.13
XSum, factKB 0.04 0.04 0.04 0.04 0.04 0.04 0.04 0.04 0.04 0.04 0.04 0.18 0.34 0.41 0.41 0.34 0.48 0.47 0.57 0.57 0.14 0.32 0.23 0.25
XSum, BERT-P 0.39 0.38 0.43 0.41 0.40 0.40 0.40 0.39 0.40 0.39 0.40 0.00 0.44 0.61 0.65 0.74 0.68 0.71 0.65 0.63 0.43 0.57 0.60 0.58
CNN/DM, ROUGE-1 0.01 0.02 0.01 0.01 0.01 0.01 0.01 0.01 0.01 0.01 0.00 0.00 0.01 0.01 0.01 0.01 0.01 0.01 0.02 0.01 0.22 0.23 0.21 0.27 0.26 0.24 0.23 0.26 0.02 0.19 0.14 0.14
CNN/DM, ROUGE-2 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 0.00 0.00 0.00 0.09 0.09 0.08 0.11 0.11 0.09 0.09 0.10 0.00 0.07 0.04 0.04
CNN/DM, ROUGE-L 0.01 0.02 0.01 0.01 0.01 0.01 0.01 0.01 0.01 0.01 0.00 0.00 0.01 0.01 0.01 0.01 0.01 0.01 0.02 0.01 0.20 0.20 0.19 0.24 0.22 0.21 0.21 0.22 0.02 0.17 0.12 0.12
CNN/DM, factKB 0.23 0.21 0.15 0.16 0.16 0.16 0.16 0.18 0.17 0.16 0.95 0.08 0.18 0.17 0.19 0.18 0.18 0.18 0.22 0.20 0.91 0.87 0.83 0.95 0.94 0.92 0.92 0.87 0.80 0.90 0.80 0.71
CNN/DM, BERT-P 0.32 0.40 0.37 0.36 0.37 0.36 0.36 0.35 0.37 0.37 0.00 0.33 0.35 0.36 0.37 0.37 0.37 0.36 0.39 0.38 0.58 0.59 0.59 0.71 0.62 0.66 0.60 0.62 0.45 0.51 0.46 0.48
CNN/DM v2, ROUGE-1 0.18 0.22
CNN/DM v2, ROUGE-2 0.08 0.10
CNN/DM v2, ROUGE-L 0.16 0.19
CNN/DM v2, factKB 0.97 0.97
CNN/DM v2, BERT-P 0.35 0.44
XSum v2, ROUGE-1 0.17
XSum v2, ROUGE-2 0.04
XSum v2, ROUGE-L 0.12
XSum v2, factKB 0.67
XSum v2, BERT-P 0.39
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
upstage/SOLAR-10.7B-Instruct-v1.0
DiscoResearch/mixtral-7b-8expert
mistralai/Mistral-7B-Instruct-v0.1
NousResearch/Yarn-Mistral-7b-128k
HuggingFaceH4/zephyr-7b-beta
NousResearch/Nous-Hermes-Llama2-13b
teknium/OpenHermes-2-Mistral-7B
HuggingFaceH4/zephyr-7b-alpha
Open-Orca/Mistral-7B-OpenOrca
ehartford/dolphin-2.1-mistral-7b
EleutherAI/gpt-neo-2.7B
EleutherAI/gpt-neo-1.3B
tiiuae/falcon-7b-instruct
togethercomputer/GPT-JT-6B-v1
EleutherAI/gpt-j-6b
KoboldAI/OPT-13B-Erebus
bigscience/bloom-7b1
pankajmathur/orca_mini_3b
ai-forever/mGPT
tiiuae/falcon-7b
bigscience/bloom-560m
EleutherAI/gpt-neo-125m