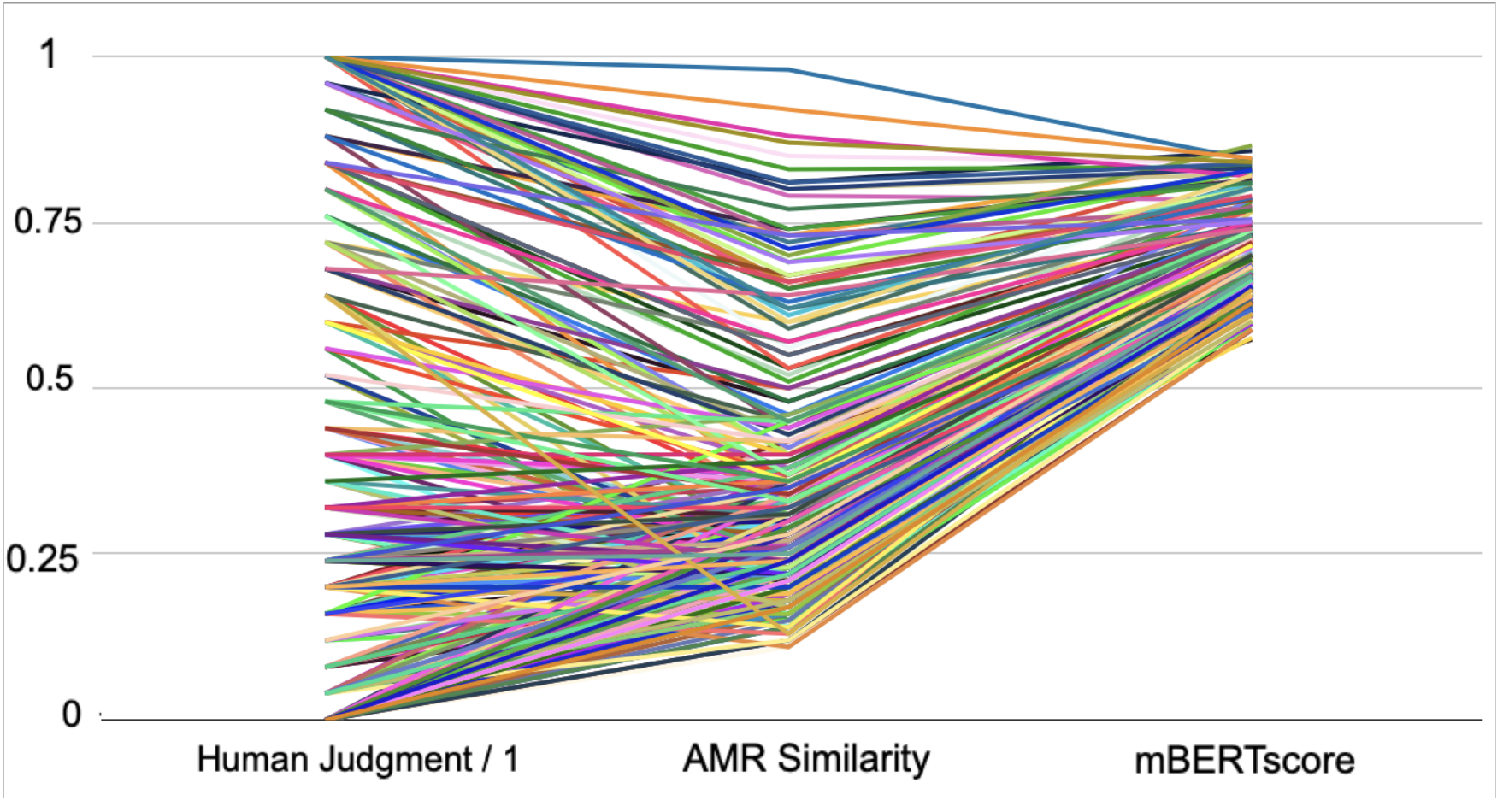
1
0.75
0.5
0.25
0
Human Judgment / 1
AMR Similarity
mBERTscore