## Normalized confusion matrix (mistral-small-latest) action -0.27 0.01 0.01 0.03 0.02 0.01 0.17 0.14 0.00 0.06 0.18 0.02 0.07 aggression -0.12 0.44 0.00 0.01 0.00 0.00 0.17 0.09 0.01 0.05 0.09 0.00 0.00 delivery -0.09 0.01 0.03 0.43 0.01 0.01 0.08 0.03 0.08 0.13 0.01 0.00 0.08 entrance -0.03 0.00 0.00 0.01 0.77 0.02 0.00 0.03 0.04 0.09 0.01 0.02 0.00 exit -0.01 0.00 0.00 0.00 0.10 0.79 0.01 0.06 0.00 0.02 0.00 0.00 0.00 True label interaction -0.10 0.02 0.00 0.00 0.00 0.00 <mark>0.70</mark> 0.05 0.00 0.01 0.13 0.00 0.00 movement -0.07 0.00 0.00 0.01 0.05 0.12 0.14 0.52 0.00 0.05 0.01 0.00 0.03 narration -0.10 0.04 0.00 0.00 0.00 0.03 0.05 0.04 0.03 0.58 0.06 0.06 0.01 object -0.06 0.02 0.01 0.00 0.01 0.00 0.12 0.01 0.00 0.03 0.70 0.02 0.00 setting -0.01 0.00 0.00 0.00 0.01 0.01 0.03 0.09 0.01 0.03 0.02 0.79 0.01 toward -0.00 0.00 0.01 0.01 0.00 0.00 0.01 0.00 0.00 0.00 0.00 0.00 0.96 actionession movement interaction MUSIC narration apaite delivery Predicted label