Normalized confusion matrix (mistral-small-latest) action -0.44 0.02 0.02 0.03 0.01 0.01 0.10 0.09 0.01 0.05 0.13 0.01 0.09 aggression -0.17 0.53 0.00 0.00 0.00 0.00 0.13 0.03 0.01 0.04 0.08 0.00 0.00 delivery -0.17 0.03 0.04 0.42 0.00 0.02 0.06 0.06 0.03 0.11 0.00 0.01 0.06 entrance -0.05 0.01 0.00 0.00 0.73 0.05 0.01 0.02 0.02 0.06 0.00 0.05 0.00 exit -0.04 0.00 0.00 0.00 0.08 0.77 0.01 0.06 0.00 0.02 0.01 0.00 0.00 True label interaction -0.17 0.03 0.00 0.00 0.00 0.00 <mark>0.68</mark> 0.04 0.00 0.01 0.06 0.00 0.02 movement -0.11 0.00 0.00 0.01 0.07 0.12 0.12 0.48 0.00 0.03 0.02 0.00 0.06 music +0.02 0.00 0.00 0.02 0.01 0.00 0.00 0.01 0.84 0.04 0.02 0.02 0.01 narration -0.14 0.07 0.00 0.01 0.00 0.02 0.05 0.03 0.03 0.59 0.02 0.03 0.01 setting -0.08 0.00 0.00 0.01 0.01 0.02 0.01 0.06 0.01 0.03 0.01 0.77 0.01 toward $+0.01\ 0.00\ 0.03\ 0.01\ 0.00\ 0.01\ 0.01\ 0.00\ 0.00\ 0.00\ 0.00\ 0.01\ 0.92$ action adjoestion movement interaction MUSIC narration apaite Predicted label