## Normalized confusion matrix (mistral-small-latest) action -0.27 0.02 0.02 0.04 0.02 0.01 0.16 0.14 0.01 0.05 0.17 0.01 0.09 aggression -0.13 0.49 0.00 0.00 0.00 0.00 0.13 0.05 0.04 0.04 0.08 0.00 0.03 delivery -0.09 0.02 0.05 0.41 0.01 0.01 0.07 0.07 0.06 0.11 0.04 0.01 0.05 entrance -0.01 0.02 0.02 0.01 0.70 0.02 0.02 0.04 0.02 0.08 0.02 0.03 0.02 exit -0.02 0.01 0.00 0.00 0.10 0.72 0.01 0.06 0.01 0.03 0.01 0.02 0.00 True label interaction -0.09 0.05 0.00 0.01 0.02 0.00 <mark>0.62</mark> 0.01 0.01 0.01 0.15 0.01 0.03 movement -0.10 0.02 0.00 0.03 0.07 0.12 0.12 0.43 0.02 0.02 0.03 0.02 0.05 music +0.03 0.01 0.01 0.01 0.01 0.01 0.01 0.02 0.78 0.05 0.04 0.02 0.01 narration -0.10 0.03 0.00 0.00 0.03 0.03 0.04 0.05 0.05 0.54 0.07 0.06 0.01 object -0.08 0.01 0.01 0.00 0.02 0.00 0.16 0.01 0.01 0.02 0.61 0.04 0.01 setting -0.02 0.02 0.01 0.01 0.03 0.02 0.01 0.10 0.01 0.05 0.02 0.71 0.01 action movement interaction narration MUSIC apaite delivery Predicted label