## Normalized confusion matrix (mistral-small-latest) action -0.49 0.01 0.00 0.01 0.06 0.00 0.04 0.25 0.01 0.01 0.07 0.04 0.00 aggression -0.16 0.67 0.00 0.00 0.00 0.03 0.08 0.00 0.01 0.05 0.00 0.00 delivery -0.20 0.01 0.00 0.42 0.02 0.01 0.05 0.11 0.05 0.02 0.00 0.10 0.00 entrance -0.01 0.00 0.01 0.01 0.83 0.03 0.01 0.04 0.01 0.01 0.01 0.05 0.00 exit +0.02 0.00 0.00 0.00 0.16 0.67 0.02 0.12 0.00 0.00 0.00 0.01 0.00 True label interaction -0.20 0.09 0.00 0.00 0.01 0.01 <mark>0.60</mark> 0.08 0.00 0.00 0.02 0.00 0.00 movement -0.06 0.00 0.00 0.00 0.05 0.11 0.03 0.74 0.00 0.00 0.00 0.01 0.00 narration -0.07 0.11 0.00 0.01 0.06 0.03 0.03 0.07 0.04 0.38 0.05 0.17 0.00 object -0.23 0.03 0.00 0.00 0.01 0.00 0.04 0.07 0.01 0.00 0.56 0.03 0.00 actionession movement interaction narration MUSIC Predicted label