## Normalized confusion matrix (mistral-small-latest) action -0.09 0.01 0.00 0.01 0.09 0.01 0.08 0.41 0.02 0.01 0.21 0.04 0.01 aggression -0.00 0.72 0.00 0.00 0.00 0.03 0.07 0.00 0.00 0.17 0.01 0.00 delivery -0.04 0.02 0.04 0.35 0.04 0.01 0.07 0.27 0.13 0.01 0.00 0.03 0.00 entrance -0.00 0.00 0.00 0.00 0.92 0.04 0.00 0.01 0.01 0.00 0.00 0.02 0.00 True label interaction -0.00 0.11 0.00 0.00 0.03 0.01 0.60 0.12 0.00 0.00 0.14 0.00 0.00 movement -0.01 0.01 0.01 0.00 0.04 0.18 0.02 0.71 0.00 0.00 0.02 0.01 0.00 music \(\frac{1}{0.00}\) 0.00 0.00 0.00 0.01 0.00 0.00 0.03 0.86 0.01 0.03 0.05 0.01 narration +0.00 0.10 0.01 0.01 0.06 0.03 0.07 0.12 0.09 0.21 0.14 0.16 0.00 object -0.00 0.02 0.00 0.00 0.01 0.00 0.02 0.05 0.01 0.00 0.86 0.02 0.00 setting -0.00 0.00 0.01 0.00 0.03 0.02 0.01 0.11 0.00 0.00 0.02 0.83 0.00 toward +0.00 0.00 0.01 0.01 0.02 0.00 0.03 0.03 0.00 0.00 0.01 0.18 0.71 action deliverytance movement interaction narration Music apaite Predicted label