Normalized confusion matrix (mistral-small-latest) action -0.12 0.03 0.03 0.02 0.05 0.06 0.15 0.11 0.03 0.04 0.12 0.10 0.14 aparte $-0.00\ 0.00\ 0.50\ 0.50\ 0.00\ 0.00\ 0.00\ 0.00\ 0.00\ 0.00\ 0.00\ 0.00$ delivery -0.00 0.02 0.02 0.53 0.02 0.02 0.11 0.03 0.00 0.14 0.02 0.06 0.05 entrance -0.00 0.00 0.00 0.00 0.58 0.13 0.05 0.00 0.00 0.00 0.03 0.21 0.00 exit -0.01 0.00 0.00 0.00 0.11 0.64 0.07 0.12 0.00 0.01 0.00 0.03 0.00 True label movement -0.11 0.06 0.08 0.11 0.11 0.08 0.03 0.00 0.03 0.06 0.08 0.11 0.14 music +0.02 0.00 0.02 0.05 0.02 0.01 0.03 0.02 0.66 0.07 0.06 0.02 0.02 narration -0.00 0.00 0.00 0.03 0.03 0.00 0.11 0.14 0.14 0.25 0.06 0.25 0.00 object -0.00 0.00 0.00 0.00 0.00 0.00 0.39 0.02 0.00 0.02 0.58 0.00 0.00 setting -0.00 0.00 0.00 0.00 0.00 0.00 0.02 0.23 0.00 0.04 0.02 0.67 0.04 actionession deliverytance movement interaction MUSIC narration apaite Predicted label