Normalized confusion matrix (mistral-small-latest) action -0.23 0.05 0.01 0.03 0.08 0.01 0.14 0.13 0.01 0.00 0.23 0.05 0.03 delivery -0.12 0.08 0.00 0.39 0.03 0.02 0.09 0.08 0.12 0.03 0.00 0.02 0.02 entrance -0.00 0.00 0.00 0.00 0.87 0.05 0.00 0.03 0.00 0.00 0.00 0.05 0.00 exit -0.01 0.01 0.00 0.00 0.14 0.77 0.00 0.05 0.00 0.00 0.00 0.01 0.00 True label interaction -0.00 0.13 0.00 0.00 0.03 0.03 <mark>0.81</mark> 0.00 0.00 0.00 0.00 0.00 0.00 movement -0.03 0.00 0.03 0.00 0.03 0.17 0.08 0.53 0.00 0.00 0.03 0.00 0.11 music +0.02 0.05 0.05 0.05 0.03 0.01 0.02 0.01 0.49 0.03 0.05 0.12 0.10 narration -0.00 0.25 0.00 0.00 0.00 0.00 0.14 0.06 0.08 0.14 0.17 0.17 0.00 object -0.00 0.02 0.00 0.03 0.00 0.00 0.10 0.02 0.00 0.00 0.82 0.02 0.00 setting -0.00 0.00 0.02 0.00 0.00 0.04 0.02 0.16 0.00 0.00 0.02 0.75 0.00 toward -0.00 0.01 0.01 0.01 0.02 0.00 0.07 0.02 0.00 0.00 0.01 0.12 0.73 actionession deliverytance interaction narration Music apaite Predicted label