Normalized confusion matrix (mistral-small-latest) action -0.22 0.05 0.01 0.03 0.08 0.01 0.14 0.14 0.01 0.01 0.22 0.05 0.03 delivery -0.12 0.08 0.00 0.38 0.03 0.02 0.11 0.08 0.14 0.02 0.00 0.03 0.00 entrance -0.00 0.00 0.00 0.00 0.89 0.05 0.00 0.00 0.00 0.00 0.00 0.05 0.00 exit -0.01 0.01 0.00 0.00 0.15 0.73 0.00 0.10 0.00 0.00 0.00 0.00 0.00 True label interaction -0.00 0.19 0.00 0.00 0.03 0.03 <mark>0.74</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 0.56 0.00 0.00 0.03 0.00 0.08 music -0.02 0.06 0.04 0.04 0.03 0.01 0.01 0.01 0.49 0.03 0.05 0.13 0.09 narration +0.00 0.19 0.00 0.00 0.00 0.00 0.14 0.08 0.08 0.14 0.19 0.17 0.00 object -0.00 0.02 0.00 0.03 0.00 0.00 0.11 0.02 0.00 0.00 0.81 0.02 0.00 setting -0.00 0.00 0.02 0.00 0.00 0.04 0.02 0.18 0.00 0.00 0.00 0.75 0.00 toward -0.00 0.01 0.01 0.01 0.03 0.00 0.07 0.02 0.00 0.00 0.01 0.11 0.73 action adjusted in deliverytance interaction narration MUSIC apaite Predicted label