Multi-Model Performance for Best, Median, and
Worst Single-Model Emotions