Experiment Comparison: progression_assessment_discussed (tb_v1_1_tb_classification_tasks)

Model
- medgemma-4b-it
- InternVL3_5-8B
- medgemma-1.5-4b-it
- OctoMed-7B
- Qwen3-VL-8B-Instruct