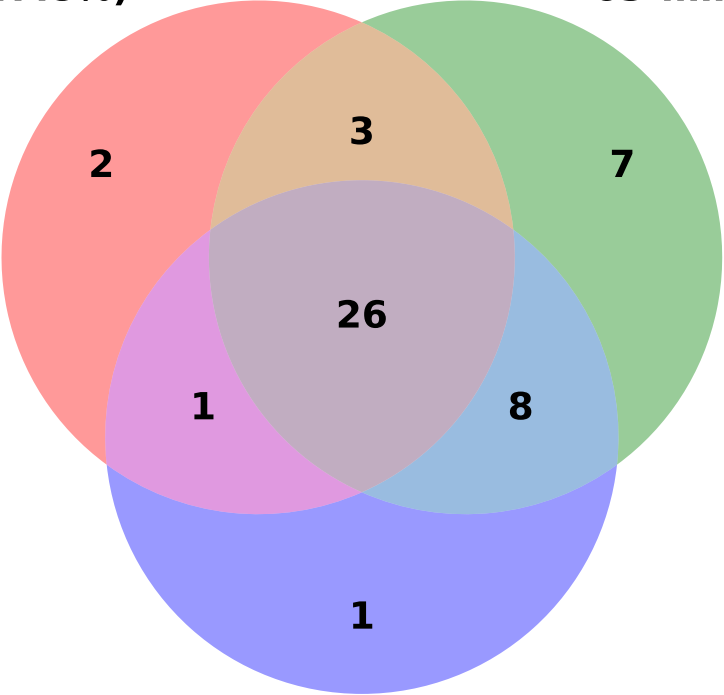


Claude 3 Sonnet (32/67, 47.8%)

o3-mini (44/67, 65.7%)



Deepseek (36/67, 53.7%)