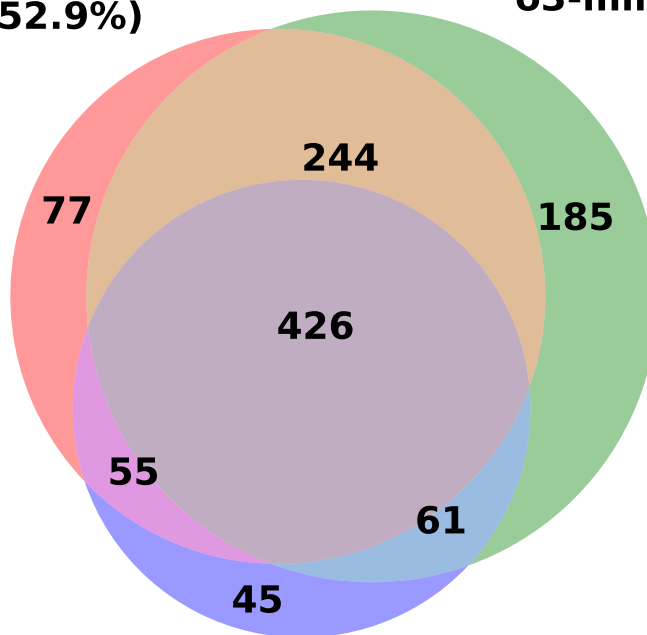


ude 3 Sonnet (802/1515, 52.9%)

o3-mini (916/1515, 60.5%)



Deepseek (587/1515, 38.7%)