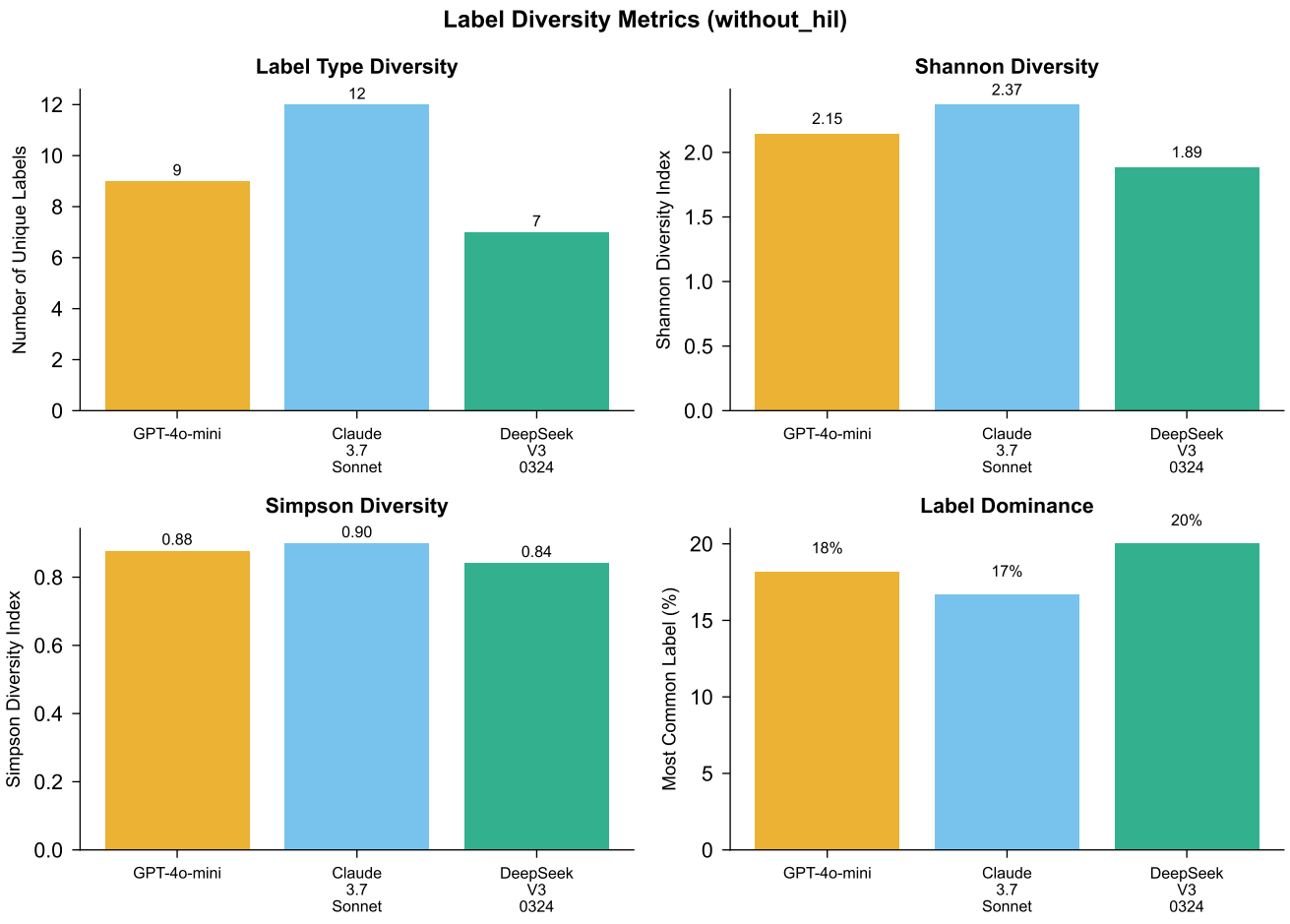
Label Diversity Metrics (without_hil)
Label Type Diversity
Number of Unique Labels
0
2
4
6
8
10
12
9
12
7
GPT-4o-mini
Claude 3.7 Sonnet
DeepSeek V3 0324
Shannon Diversity
Shannon Diversity Index
0.0
0.5
1.0
1.5
2.0
2.37
2.15
1.89
GPT-4o-mini
Claude 3.7 Sonnet
DeepSeek V3 0324
Simpson Diversity
Simpson Diversity Index
0.0
0.2
0.4
0.6
0.8
0.88
0.90
0.84
GPT-4o-mini
Claude 3.7 Sonnet
DeepSeek V3 0324
Label Dominance
Most Common Label (%)
0
5
10
15
20
18%
17%
20%
GPT-4o-mini
Claude 3.7 Sonnet
DeepSeek V3 0324