

Correlation with True GE vs Retrieval Performance ( $-\log(\text{MRR})$ )

