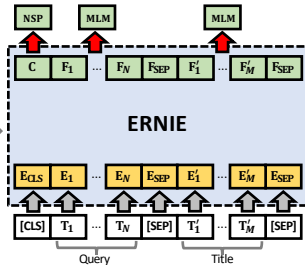
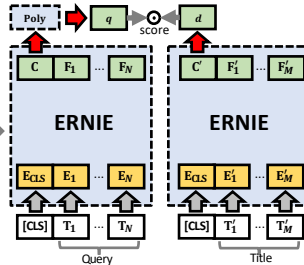


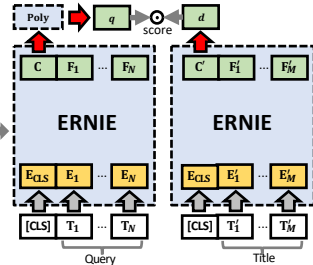
Stage 1: pretraining
Model: n -layer Transformer
Data: trillion-scale Heterogeneous corpus (Chinese Wikipedia, Baidu news & Baike & Tieba)
Tasks: entity- and phrase-level MLM; NSP
Training: unsupervised



Stage 2: post-pretraining
Model: n -layer Transformer
Data: trillion-scale search logs
Tasks: : entity- and phrase-level MLM; NSP
Training: unsupervised



Stage 3: intermediate fine-tuning
Model: the retrieval model
Data: trillion-scale search logs
Tasks: semantic matching
Training: supervised



Stage 4: target fine-tuning
Model: the retrieval model
Data: small-scale manually labeled data
Tasks: semantic matching
Training: supervised