TinyBert

Attention head: 02

Layer: 02

Train : Test ratio = 0.8 : 0.2

Dataset	Evaluation Loss	Perplexity
Full Dataset	2.245218	9.4425
Sadhu	2.2859903	9.8354
Social Media	2.3541173	10.5288
Newspaper	2.2503764	9.4913