

	SQuAD1.1	SQuAD2.0	MNLI	SST-2	RACE	Avg
No additional data	89.3 /82.3	80.0 /77.1	81.6	90.3	64.0	80.1
With additional data	88.8/81.7	79.1/76.3	82.4	92.8	66.0	80.8

Table 7: The effect of additional training data using the ALBERT-base configuration.