Device			NVIDIA A40 (GPU)				Raspberry Pi (4 CPU)	
Training Method	Tunable Paras. (M)	Model Size ¹ (M)	Convergence (Round)	FedAvg. (Accuracy)	Convergence (Epoch)	Centralize. (Accuracy)	Latency (S)	Memory (G)
Origin	67.0	268.0	263	0.8305	4 (3.7)	0.8338	18.27	3.0
Freeze: E	43.1	172.4	151	0.8125	5 (4.9)	0.8376	17.03	2.8
Freeze: E+L ₀	36.0	144.0	161	0.8119	6 (5.1)	0.8368	15.03	2.4
Freeze: E+L ₀₋₁	29.0	116.0	176	0.8058	6 (5.1)	0.8338	13.12	2.2
Freeze: E+L ₀₋₂	21.9	87.6	174	0.8027	6 (5.3)	0.8286	11.83	1.8
Freeze: E+L ₀₋₃	14.8	59.2	610	0.8376	7 (6.7)	0.8386	9.63	1.6
Freeze: E+L ₀₋₄	7.7	30.8	850	0.8225	8 (7.4)	0.822	8.02	1.3
Freeze: E+L ₀₋₅	0.6	2.4	850	0.6563	34 (33.2)	0.7065	5.28	1.0
Adapter: Origin	1.05	4.2	717	0.8267	14 (13.3)	0.8261	12.01	2.0
A-Freeze ² : E	1.05	4.2	850	0.8226	20	0.8347	11.65	2.0
A-Freeze: E+L ₀	0.98	3.92	850	0.8232	20	0.834	10.78	1.8
A-Freeze: E+L ₀₋₁	0.90	3.6	850	0.8149	20	0.832	9.87	1.4
A-Freeze: E+L ₀₋₂	0.83	3.32	850	0.8131	20	0.8212	8.76	1.5
A-Freeze: E+L ₀₋₃	0.76	3.04	850	0.8039	20	0.8086	7.85	1.1
A-Freeze: E+L ₀₋₄	0.68	2.72	850	0.7003	20	0.7084	6.98	1.1
A-Freeze: E+L ₀₋₅	0.61	2.44	850	0.6563	20	0.6997	7.05	1.0

^{1.} Tunable Model Size = Tunable Parameters * 4. Original DistilBert Size is 256MB. Bert is 414MB in HDF5.

^{2.} Adapter + Layer-Freeze. Adapter size is 48.