

In general, the best time per iteration follows closely the optimal time to convergence. Parallel CPU is the fastest on sparse data for the same reasons mentioned previously. With the exception of SVM on **w8a**, GPU has the fastest time per iteration on dense data. This proves that – with the optimal data layout – the superior FLOPS on the GPU lead to better performance. Identifying the optimal layout, however, requires the in-depth analysis performed in this work. The higher computational complexity of LR is reflected in the higher time per iteration across all the configurations. Table ?? summarizes the speedup in time per iteration corresponding to these values. In the best case, parallel CPU achieves a speedup of 6X over sequential CPU on **news** which is consistent with results published in the literature . The best speedup of GPU over parallel CPU is at most 5X on **covtype**.

task	dataset	cpu-seq/cpu-par	cpu-par/gpu
LR	covtype	0.60	5.80
	w8a	2.54	2.11
	real-sim	3.09	0.30
	rcv1	4.86	0.31
	news	6.09	0.13
SVM	covtype	0.69	5.13
	w8a	0.39	2.15
	real-sim	1.45	0.54
	rcv1	3.18	0.72
	news	5.60	0.17