

MODEL	model_type	data_size	seq_length	batch_size	epochs	Adam: lr	Adam: epsilon	Warmup_prop	max_norm	accum	met_sch	accum	crit	TR Acc	VAL Acc	TR loss	VAL loss	time/epoch
BASE	albert	Total	70	-	-	-	-	-	-	-	-	-	mean	-	57,23%	-	0,6898	
BASE	bart	Total	70	-	-	-	-	-	-	-	-	-	mean	-	-	-	-	
BASE	bert	Total	70	-	-	-	-	-	-	-	-	-	mean	-	60,96%	-	0,7709	
BASE	camembert	Total	70	-	-	-	-	-	-	-	-	-	mean	-	62,89%	-	0,6796	
BASE	distilbert	Total	70	-	-	-	-	-	-	-	-	-	mean	-	37,10%	-	0,7175	
BASE	flaubert	Total	70	-	-	-	-	-	-	-	-	-	mean	-	43,16%	-	0,8129	
BASE	roberta	Total	70	-	-	-	-	-	-	-	-	-	mean	-	37,10%	-	0,7108	
BASE	xlm	Total	70	-	-	-	-	-	-	-	-	-	mean	-	42,17%	-	1,2632	
BASE	xlm_roberta	Total	70	-	-	-	-	-	-	-	-	-	mean	-	62,90%	-	0,6623	
BASE	xl_net	Total	70	-	-	-	-	-	-	-	-	-	mean	-	37,10%	-	0,8550	
BASE	DistilBERT	Total	60	-	-	-	-	-	-	-	-	-	mean	-	37,10%	-	0,7176	
1	DistilBERT	125k	60	8	2	1,00E-03	1,00E-08	0,01/accum	1	no	lin_antic	16	sum	62,67%	63,12%	0,6471	0,6427	
2	DistilBERT	125k	70	8	2	1,00E-03	1,00E-08	0,01/accum	1	no	lin_antic	16	sum	62,70%	63,12%	0,6477	0,6427	
3	DistilBERT	125k	80	8	2	1,00E-03	1,00E-08	0,01/accum	1	no	lin_antic	16	sum	62,69%	63,12%	0,6479	0,6427	
4	DistilBERT	125k	90	8	2	1,00E-03	1,00E-08	0,01/accum	1	no	lin_antic	16	sum	62,71%	63,12%	0,6479	0,6427	
6	DistilBERT	Total	60	8	2 (3 of)	2,00E-05	1,00E-08	0,01/accum	1	no	const	16	sum	90,22%	89,05%	0,3843	0,4586	
7	DistilBERT	Total	60	16	2	2,00E-05	1,00E-08	0,01/accum	1	no	const	16	sum	90,78%	89,14%	0,2343	0,2690	
8	DistilBERT	Total	60	32	2	2,00E-05	1,00E-08	0,01/accum	1	no	const	16	sum	90,61%	88,82%	0,2260	0,2757	19
9	DistilBERT	Total	60	16	2	2,00E-05	1,00E-08	0,01/accum	1	no	const	16	mean	90,68%	89,00%	0,2331	0,2800	
10	DistilBERT	Total	60	16	2	2,00E-05	1,00E-08	0,01	1	no	const	-	sum	90,68%	89,52%	0,2358	0,2799	
11	DistilBERT	Total	60	32	2 (of 3)	1,00E-05	1,00E-08	0,01	1	no	const	-	mean	89,33%	88,92%	0,2504	0,2568	19
12	DistilBERT	Total	60	32	4 (of 5)	2,00E-06	1,00E-08	0,01	1	no	const	-	mean	88,39%	88,02%	0,2678	0,2825	12
13	DistilBERT	Total	60	32	2	2,00E-05	1,00E-08	0,0001	1	no	const	-	mean	90,66%	89,18%	0,2251	0,2593	19
14	DistilBERT	Total	60	128	2	2,00E-05	1,00E-08	0,01	1	no	const	-	mean	89,54%	89,25%	0,2435	0,2520	15
15	DistilBERT	Total	60	128	2	2,00E-05	1,00E-08	0,01	1	no	lin	-	mean	89,11%	88,59%	0,2516	0,2628	16
16	DistilBERT	Total	60	128	2	2,00E-05	1,00E-08	0,01	1	no	no	-	mean	89,46%	88,96%	0,2462	0,2617	16
17	DistilBERT	Total	60	64	2	2,00E-05	1,00E-08	0,01	1	no	const	-	mean	90,29%	89,13%	0,2303	0,2514	17
18	DistilBERT	Total	60	256	2	2,00E-05	1,00E-08	0,01	1	no	const	-	mean	88,73%	88,53%	0,2603	0,2636	15
19	DistilBERT	Total	60	512	2	2,00E-05	1,00E-08	0,01	1	no	const	-	mean	87,70%	87,93%	0,2790	0,2759	14
20	DistilBERT	Total	70	128	2	2,00E-05	1,00E-08	0,01	1	no	const	-	mean	89,51%	89,36%	0,2447	0,2499	18
21	DistilBERT	Total	70	128	4	5,00E-06	1,00E-08	0,01	1	no	const	-	mean	89,12%	88,23%	0,2521	0,2728	18
22	DistilBERT	Total	70	128	2	1,00E-05	1,00E-08	0,01	1	no	const	-	mean	88,02%	88,05%	0,2736	0,2717	18
23	DistilBERT	Total	70	128	2	2,00E-05	1,00E-08	0,01	1	no	const	-	sum	89,49%	89,30%	0,2442	0,2516	18
24	DistilBERT	Total	70	128	2	2,00E-05	1,00E-08	0,01	2	no	const	-	mean	89,62%	89,18%	0,2552	0,2424	18
25	DistilBERT	Total	70	128	2	2,00E-05	1,00E-08	0,01	0,5	no	const	-	mean	89,57%	89,18%	0,2432	0,2557	18
26	DistilBERT	Total	70	128	2	2,00E-05	1,00E-08	0,01	5	no	const	-	mean	89,55%	89,18%	0,2435	0,2570	17
27	DistilBERT	Total	70	128	2	2,00E-05	1,00E-08	0,001	1	no	const	-	mean	89,43%	89,20%	0,2540	0,2456	18
28	DistilBERT	Total	70	128	2	2,00E-05	1,00E-08	0,01	0	no	const	-	mean	61,40%	62,94%	0,6793	0,6383	18
XLM_ROBERTA																		
BASE	xlm_roberta	Total	70	-	-	-	-	-	-	-	-	-	mean	-	62,90%	-	0,6623	
1	xlm_roberta	Total	60	4	2	2,00E-05	1,00E-08	0,01	1	no	const	-	mean	52,79%	62,90%	140368,88	54258,79	1209
2	xlm_roberta	90k	60	6	2	2,00E-05	1,00E-08	0,01	1	no	const	-	mean	84,76%	83,89%	0,4314	0,5062	22
3	xlm_roberta	90k	60	8	2	2,00E-05	1,00E-08	0,01	1	no	const	-	mean	62,30%	62,88%	nan	nan	
4	xlm_roberta	Total	60	6	2 (of 4)	2,00E-05	1,00E-08	0,01	1	no	const	-	mean	85,85%	86,79%	0,4071	0,4033	96
5	xlm_roberta	Total	60	64	3	2,00E-05	1,00E-08	0,01	1	no	const	-	mean	91,89%	89,98%	0,1982	0,2447	39
6	xlm_roberta	Total	70	64	2 (of 3)	2,00E-05	1,00E-08	0,001	1	no	const	-	mean	89,45%	89,53%	0,2464	0,2460	39
7	xlm_roberta	Total	60	64	3	2,00E-05	1,00E-08	0,001	1	no	const	-	mean	91,73%	90,13%	0,2009	0,2531	36
ROBERTA																		
BASE	roberta	Total	70	-	-	-	-	-	-	-	-	-	mean	-	37,10%	-	0,7108	
1	roberta	Total	70	8	2	2,00E-05	1,00E-08	0,01	1	no	const	-	mean	63,13%	62,90%	nan	nan	65
2	roberta	Total	70	6	2	2,00E-05	1,00E-08	0,01	1	no	const	-	mean	63,13%	62,90%	nan	nan	76
3	roberta	Total	70	64	2	2,00E-05	1,00E-08	0,01	1	no	const	-	mean	90,42%	90,11%	0,2296	0,2481	37
4	roberta	Total	70	64	3	2,00E-05	1,00E-08	0,01	1	no	const	-	mean	-	-	-	-	37
5	roberta	Total	70	128	2	2,00E-05	1,00E-08	0,01	1	no	const	-	mean	90,16%	89,94%	0,2339	0,2385	35
6	roberta	Total	70	64	1 (of 2)	3,00E-05	1,00E-08	0,01	1	no	const	-	mean	85,89%	89,20%	0,3155	0,2510	37
7	roberta	Total	70	64	2	2,00E-05	1,00E-08	0,001	1	no	const	-	mean	90,45%	90,27%	0,2285	0,2470	38
8	roberta	Total	70	32	2	2,00E-05	1,00E-08	0,001	1	no	const	-	mean	90,40%	90,24%	0,2340	0,2427	72
9	roberta	Total	70	64	3	8,00E-06	1,00E-08	0,001	1	no	const	-	mean	91,33%	90,30%	0,2095	0,2386	37
10	roberta	Total	70	64	3 (of 4)	7,00E-06	1,00E-08	0,001	1	no	const	-	mean	90,92%	90,30%	0,2168	0,2410	37
11	roberta	Total	70	64	3 (of 4)	5,00E-06	1,00E-08	0,001	1	no	const	-	mean	90,13%	89,93%	0,2326	0,2418	38
XLNET																		
BASE	xlnet	Total	70	-	-	-	-	-	-	-	-	-	mean	-	37,10%	-	0,8550	
1	xlnet	Total	70	64	3	2,00E-05	1,00E-08	0,01	1	no	const	-	mean	93,14%	90,44%	0,1732	0,2509	55
2	xlnet	Total	70	64	2 (of 3)	6,00E-06	1,00E-08	0,01	1	no	const	-	mean	88,60%	88,76%	0,2640	0,2644	54
3	xlnet	Total	70	128	3	2,00E-05	1,00E-08	0,01	1	no	const	-	mean	90,15%	89,89%	0,2340	0,2474	52
4	xlnet	Total	70	64	3	2,00E-05	1,00E-08	0,001	1	no	const	-	mean	93,16%	90,62%	0,1729	0,2463	54
5	xlnet	Total	70	32	3	2,00E-05	1,00E-08	0,001	1	no	const	-	mean	93,28%	90,48%	0,1771	0,2629	58
XLM																		
BASE	xlm	Total	70	-	-	-	-	-	-	-	-	-	mean	-	42,17%	-	1,2632	
1	xlm	Total	70	32	1 (of 2)	2,00E-05	1,00E-08	0,001	1	no	const	-	mean	61,22%	63,05%	0,6818	0,6813	231
BERT																		
BASE	bert	Total	70	-	-	-	-	-	-	-	-	-	mean	-	60,96%	-	0,7709	
1	bert	Total	70	16	1 (of 2)	2,00E-05	1,00E-08	0,01	1	no	const	-	sum	86,35%	88,58%	0,3127	0,2682	49
2	bert	Total	70	64	2	1,00E-05	1,00E-08	0,01	1	no	const	-	sum	90,52%	90,14%	0,2252	0,2384	37
3	bert	Total	70	64	2 (of 4)	1,00E-05	1,00E-08	0,01	1	no	const	-	mean	90,54%	90,19%	0,2260	0,2384	37
4	bert	Total	70	64	3	8,00E-06	1,00E-08	0,01	1	no	const	-	sum	92,48%	90,21%	0,1858	0,2438	38
5	bert	Total	70	64	3 (of 4)	6,00E-06	1,00E-08	0,01	1	no	const	-	mean	91,59%	89,96%	0,2044	0,2436	40
6	bert	Total	70	128	2 (of 3)	8,00E-06	1,00E-08	0,01	1	no	const	-	mean	88,39%	88,86%	0,2654	0,2583	35
7	bert	Total	70	64	3	8,00E-06	1,00E-08	0,001	1	no	const	-	mean	92,49%	90,44%	0,1861	0,2434	37
8																		

ALBERT																			
BASE	albert	Total	70	-	-	-	-	-	-	-	-	-	mean	-	57.23%			0.6898	
1	albert	Total	70	64	2 (of 3)	2,00E-05	1,00E-08	0,01	1	no	const	-	mean	91,42%	90,36%	0,2097	0,2332		127
2	albert	Total	70	64	2	2,00E-05	1,00E-08	0,001	1	no	const	-	mean	91,37%	90,06%	0,2121	0,2377		127
3	albert	Total	70	32	2	2,00E-05	1,00E-08	0,01	1	no	const	-	mean	90,69%	89,47%	0,2278	0,2517		136
4	albert	Total	70	64	2 (of 3)	8,00E-06	1,00E-08	0,01	1	no	const	-	mean	91,04%	90,10%	0,2178	0,2377		125
5	albert	Total	70	64	3	4,00E-06	1,00E-08	0,01	1	no	const	-	mean	91,96%	90,06%	0,1986	0,2446		127