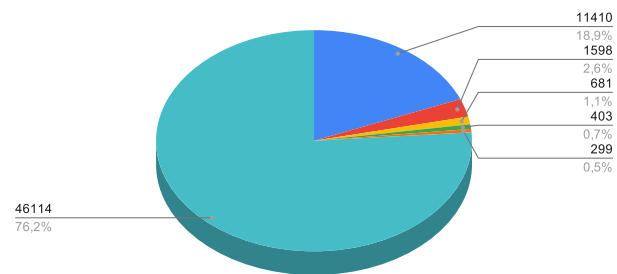Example100words   per_device_train_batch = 8   gradient_accumulation_steps=4

Model	epochs	learning rate	p@1	p@2	p@3	p@4	p@5	No acertadas
Roberta-Large	Base	Base	11	4	6	0	3	76
Roberta-Large	3	5,00E-05	14	2	5	5	1	73
Roberta-Large	5	5,00E-05	17	4	4	1	2	72
Roberta-Large	10	5,00E-05	18	8	2	0	3	69
Roberta-Large	15	5,00E-05	21	6	3	4	2	64
Roberta-Large	20	5,00E-05	22	12	3	1	0	62
Roberta-Large	30	5,00E-05	30	8	3	2	1	56
Roberta-Large	50	5,00E-05	47	3	0	0	0	50
Roberta-Large	100	5,00E-05	51	1	0	0	0	48
Roberta-Large	Base	Base	11	4	6	0	3	76
Roberta-Large	3	5,00E-06	14	2	4	1	3	76
Roberta-Large	5	5,00E-06	14	2	4	2	2	76
Roberta-Large	10	5,00E-06	14	3	4	4	2	73
Roberta-Large	15	5,00E-06	14	5	3	6	0	72
Roberta-Large	20	5,00E-06	18	3	2	4	2	71
Roberta-Large	30	5,00E-06	20	2	4	1	4	69
Roberta-Large	50	5,00E-06	21	2	6	3	4	64
Roberta-Large	100	5,00E-06	22	11	3	5	2	57
Roberta-Large	Base	Base	11	4	6	0	3	76
Roberta-Large	3	7,50E-06	13	2	5	0	3	77
Roberta-Large	5	7,50E-06	13	3	4	1	3	76
Roberta-Large	10	7,50E-06	14	2	4	4	4	72
Roberta-Large	15	7,50E-06	15	5	4	4	1	71
Roberta-Large	20	7,50E-06	19	3	3	5	1	69
Roberta-Large	30	7,50E-06	20	3	6	5	0	66
Roberta-Large	50	7,50E-06	22	5	5	4	1	63
Roberta-Large	100	7,50E-06	31	6	6	4	1	52
Roberta-Large	Base	Base	11	4	6	0	3	76
Roberta-Large	3	1,00E-05	13	2	5	1	3	76
Roberta-Large	5	1,00E-05	14	2	2	3	2	77
Roberta-Large	10	1,00E-05	13	3	5	6	1	72
Roberta-Large	15	1,00E-05	15	5	6	2	3	69
Roberta-Large	20	1,00E-05	18	3	8	3	1	67
Roberta-Large	30	1,00E-05	21	3	5	4	1	66
Roberta-Large	50	1,00E-05	21	7	4	5	1	62
Roberta-Large	100	1,00E-05	39	6	3	0	2	50
Roberta-Large	Base	Base	11	4	6	0	3	76
Roberta-Large	3	2,50E-05	12	2	3	2	3	78
Roberta-Large	5	2,50E-05	11	2	10	0	1	76
Roberta-Large	10	2,50E-05	17	5	3	4	2	69
Roberta-Large	15	2,50E-05	17	9	3	3	0	68
Roberta-Large	20	2,50E-05	18	7	6	5	0	64
Roberta-Large	30	2,50E-05	23	9	5	3	1	59
Roberta-Large	50	2,50E-05	39	4	3	2	1	51
Roberta-Large	100	2,50E-05	51	0	0	0	0	49
Roberta-Large	Base	Base	11	4	6	0	3	76
Roberta-Large	3	7,50E-05	8	5	4	2	3	78
Roberta-Large	5	7,50E-05	10	8	3	0	1	78
Roberta-Large	10	7,50E-05	20	6	3	3	2	66
Roberta-Large	15	7,50E-05	22	10	4	2	1	61

Model	p@1	p@2	p@3	p@4	p@5	No acertadas
Roberta-base	11	4	6	0	3	76
Roberta-large	16	7	8	1	1	67
xlm-roberta-base	3	0	2	3	0	92
xlm-roberta-large	1	1	0	1	0	97
Bart-base	1	2	0	1	1	95
bart-large	1	1	0	2	7	89
albert-base-v2	3	1	2	0	0	94
albert-large-v2	5	3	1	1	0	90
distilbert-base-uncased	4	2	0	1	0	93
distilbert-base-cased	2	1	3	3	0	91
deberta-xlarge-mnli	0	0	0	0	0	100
deberta-base	0	0	0	0	0	100
flaubert-base-uncased	0	0	0	0	0	100
flaubert-base-cased	0	0	0	0	0	100
camembert-base	0	0	0	0	0	100

Epoch	lr (best top-1)	lr (best top-5)
3	5,00E-05	1,00E-04
5	1,00E-04	1,00E-04
10	1,00E-04	1,00E-04
15	2,50E-04	1,00E-04
20	1,00E-04	7,50E-05
30	1,00E-04	7,50E-05
50	7,50E-05	7,50E-05
100	5,00E-05	5,00E-05

Seen

P@1	P@10	P@100	Sin acertar
30 (6%)	77 (15.4%)	108 (21.6%)	392

Unseen

20 (4%)	70 (14%)	117 (23.4%)	383

Desc

51 (25.5%)	130 (65%)	164 (82%)	36

Roberta-Large	20	7,50E-05	27	10	1	3	0	**59**
Roberta-Large	30	7,50E-05	39	3	5	0	2	**51**
Roberta-Large	50	7,50E-05	**50**	1	0	0	0	49
Roberta-Large	100	7,50E-05	51	0	0	0	0	49
Roberta-Large	Base	Base	11	4	6	0	3	76
Roberta-Large	3	1,00E-04	13	5	9	3	2	**68**
Roberta-Large	5	1,00E-04	**19**	4	3	1	1	**72**
Roberta-Large	10	1,00E-04	**22**	5	5	1	3	**64**
Roberta-Large	15	1,00E-04	23	9	4	1	2	**61**
Roberta-Large	20	1,00E-04	**31**	5	1	2	1	60
Roberta-Large	30	1,00E-04	**40**	3	4	0	0	53
Roberta-Large	50	1,00E-04	49	1	0	0	0	50
Roberta-Large	100	1,00E-04	51	0	0	0	0	49
Roberta-Large	Base	Base	11	4	6	0	3	76
Roberta-Large	3	2,50E-04	6	2	1	2	2	87
Roberta-Large	5	2,50E-04	7	7	2	3	2	79
Roberta-Large	10	2,50E-04	14	5	7	6	1	67
Roberta-Large	15	2,50E-04	**24**	5	5	0	1	65
Roberta-Large	20	2,50E-04	31	3	5	1	0	60
Roberta-Large	30	2,50E-04	41	3	3	1	0	52
Roberta-Large	50	2,50E-04	50	0	1	0	0	49
Roberta-Large	100	2,50E-04	51	0	0	0	0	49
Roberta-Large	Base	Base	11	4	6	0	3	76
Roberta-Large	3	5,00E-04	9	2	2	1	1	85
Roberta-Large	5	5,00E-04	10	5	0	2	0	83
Roberta-Large	10	5,00E-04	15	2	0	1	2	80
Roberta-Large	15	5,00E-04	22	9	1	0	0	68
Roberta-Large	20	5,00E-04	25	6	1	3	3	62
Roberta-Large	30	5,00E-04	34	6	5	1	2	52
Roberta-Large	50	5,00E-04	44	4	2	1	0	49
Roberta-Large	100	5,00E-04	51	0	0	0	0	49

WORDNET

	Finetuned with	epochs	learning rate	p@1	p@2	p@3	p@4	p@5	No acertadas		
Roberta-large	Base	Base	Base	5556	1853	1021	709	494	50872	60505	9633
Roberta-large	Wordnet	10	1,00E-04	11410	1598	681	403	299	46114		14391

LDOCE 1000 WORDS

	Finetuned with	epochs	learning rate	p@1	p@2	p@3	p@4	p@5	No acertadas
Roberta-large	Base	Base	Base	110	40	19	13	13	805
Roberta-large	WordNet	10	1,00E-04	26	15	5	6	8	940

DEF 1000 WORDS

	Finetuned with	epochs	learning rate	p@1	p@2	p@3	p@4	p@5	No acertadas
Roberta-large	Base	Base	Base	265	53	17	15	8	642
Roberta-large	WordNet	10	1,00E-04	116	35	15	14	8	812

Finetuned roberta-large

11410
18,9%
1598
2,6%
681
1,1%
403
0,7%
299
0,5%

46114
76,2%

Roberta-large

5556
9,2%
1853
3,1%
1021
1,7%
709
1,2%

50872
84,1%